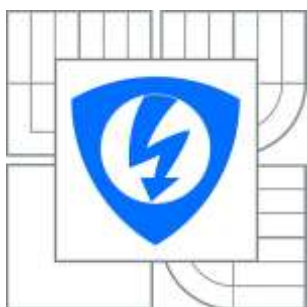




VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH
TECHNOLOGIÍ
ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF BIOMEDICAL ENGINEERING

ANALÝZA CYTOLOGICKÝCH SNÍMKŮ

ANALYSIS OF CYTOLOGY IMAGES

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. JAN PAVLÍK

VEDOUCÍ PRÁCE

SUPERVISOR

doc. Ing. RADIM KOLÁŘ, Ph.D.

BRNO 2012



VYSOKÉ UČENÍ
TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

Ústav biomedicínského inženýrství

Diplomová práce

magisterský navazující studijní obor
Biomedicínské inženýrství a bioinformatika

Student: Bc. Jan Pavlík
Ročník: 2

ID: 106123
Akademický rok: 2011/2012

NÁZEV TÉMATU:

Analýza cytologických snímků

POKYNY PRO VYPRACOVÁNÍ:

1) Proveďte literární rešerši na téma digitální analýzy cytologických snímků se zaměřením na klasifikaci základních typů bílých krvinek. 2) Dále nasnímejte sadu cytologických snímků ve školní laboratoři, kterou použijete v práci. 3) V další fázi navrhnete postup pro automatickou klasifikaci bílých krvinek - zabývejte se předzpracováním snímků, segmentací, extrakcí příznaků a samotnou klasifikací. 4) Za tímto účelem nastudujte potřebné postupy z oblasti zpracování obrazů. 5) Navržený postup realizujte v programovém prostředí Matlab a vyhodnoťte jeho úspěšnost z hlediska chyby klasifikace na nasnímané sadě obrazů.

DOPORUČENÁ LITERATURA:

- [1] COLANTONIO, S a kol. Cell image analysis ontology, Pattern Recognition and Image Analysis 18 (2), 332-341, 2007
[2] THEERA-UMPON, N.; DHOMPONGSA, S. Morphological Granulometric Features of Nucleus in Automatic Bone Marrow White Blood Cell Classification, IEEE Transactions on Information Technology in Biomedicine, 11 (3), 353-359, 2007

Termín zadání: 6.2.2012

Termín odevzdání: 18.5.2012

Vedoucí práce: doc. Ing. Radim Kolář, Ph.D.

Konzultanti diplomové práce:

prof. Ing. Ivo Provazník, Ph.D.

Předseda oborové rady

UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

Abstrakt

Tato diplomová práce je zaměřena na automatizaci procesu diferenciálního rozpočtu leukocytů v periferní krvi založeném na zpracování obrazů. Zabývá se návrhem celého procesu zpracování digitálních snímků – od snímání a předzpracování snímků, segmentace jádra a cytoplazmy, výběrem příznaků a klasifikátoru, včetně jeho testování na sadě obrazů, která byla nasnímána také v rámci této práce. V teoretické části práce jsou popsány dostupné segmentační metody a klasifikační postupy. Tato práce se v praktické části zabývá segmentačními procesy, které mají za úkol separovat jádro a cytoplazmu leukocytů. Na základě těchto vyhrazených struktur je provedena jejich statistická analýza. Podle reprezentativních statistických parametrů je vybrán soubor příznaků. Tato data pak vstupují do klasifikačního procesu, realizovaného třemi umělými neuronovými sítěmi. Celkem bylo hodnoceno 5 tříd leukocytů: neutrofilní granulocyty, lymfocyty, monocyty, eozinofilní a bazofilní granulocyty. Senzitivita a specifita klasifikace u 4 z 5 (neutrofilní granulocyty, lymfocyty, monocyty, eozinofilní granulocyty) buněčných populací je vyšší než 90 %. Senzitivita klasifikace bazofilních granulocytů byla vyhodnocena na 75 % a specifita na 67 %. Celková schopnost klasifikace byla otestovaná na 111 buňkách a vykazovala přibližně 91% úspěšnost. Všechny algoritmy jsou realizovány v prostředí MATLAB.

Klíčová slova

cytologie, leukocyty, diferenciální rozpočet leukocytů v periferní krvi, krevní nátěr, digitální morfologie, barevné prostory, prahování, segmentace, morfologické operace, filtrace, metoda gradient vector flow, watershed klasifikace, umělé neuronové sítě, metoda support vector machine, narůstání regionů MATLAB.

Abstract

This master's thesis is focused on automating the process of differential leukocyte count in peripheral blood using image processing. It deals with the design of the processing of digital images - from scanning and image preprocessing, segmentation nucleus and cytoplasm, feature selection and classifier, including testing on a set of images that were scanned in the context of this work. This work introduces used segmentation methods and classification procedures which separate nucleus and the cytoplasm of leukocytes. A statistical analysis is performed on the basis of these structures. Following adequate statistical parameters, a set of features has been chosen. This data then go through a classification process realized by three artificial neural networks. Overall were classified 5 types of leukocytes: neutrophils, lymphocytes, monocytes, eosinophiles and basophiles. The sensitivity and specificity of the classification made for 4 out of 5 leukocyte types (neutrophils, lymphocytes, monocytes, eosinophiles) is higher than 90 %. Sensitivity of classification basophiles was evaluated at 75 % and specificity at 67 %. The total ability of classification has been tested on 111 leukocytes and was approximately 91% successful. All algorithms were created in the MATLAB program.

Keywords

cytology, leukocytes, differential leukocyte count in peripheral blood, blood smear, digital morphology, colour spaces, thresholding, segmentation, mathematical morphology, filtering, the method of gradient vector flow, Watershed classification, artificial neural network, support vector machine method, region growing, MATLAB.

Bibliografická citace práce

PAVLÍK, J. *Analýza cytologických snímků*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2012. 76 stran. Vedoucí diplomové práce doc. Ing. Radim Kolář, Ph.D.

Prohlášení

Prohlašuji, že svoji diplomovou práci na téma Analýza cytologických snímků jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené diplomové práce dále prohlašuji, že v souvislosti s vytvořením tohoto projektu jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

V Brně dne 16. kvěna 2012

.....

podpis autora

Poděkování

Děkuji mé rodině, blízkým přátelům a kolegům za jejich podporu v době studia. Děkuji vedoucímu diplomové práce doc. Ing. Radimovi Kolářovi, Ph.D. za odbornou pomoc a výborný přístup. Děkuji panu doc. RNDr. Jaromírovi Baštincovi, CSc. za pomoc při statistické analýze této práce, jeho osobní podporu, nadhled a cenné rady. Poděkování také patří RNDr. Janě Šmídové za půjčení odborných materiálů firmy SYSMEX CZ s.r.o.

V Brně dne 16. kvěna 2012

.....

podpis autora

Obsah

1	Úvod.....	12
2	Cíl práce	13
3	Hematologický základ	14
3.1	Charakteristika leukocytů	14
3.2	Třídy leukocytů.....	15
3.3	Leukocyty hodnocené v barveném nátěru periferní krve	15
3.3.1	Nesegmentovaný granulocyt (tyč)	15
3.3.2	Segmentovaný neutrofilní granulocyt (neutrofilní segment, neutrofil)	15
3.3.3	Segmentovaný eozinofilní granulocyt (eozinofilní segment, eozinofil) ...	16
3.3.4	Segmentovaný bazofilní granulocyt (bazofilní segment, bazofil)	16
3.3.5	Monocyt	17
3.3.6	Malý lymfocyt.....	17
3.3.7	Velký lymfocyt	18
3.4	Procentuální zastoupení leukocytů v periferní krvi	18
4	Literární průzkum programového zpracování	19
4.1	Využití barevných prostorů	19
4.2	Předzpracování	19
4.3	Segmentace	19
4.3.1	Segmentace matematickými morfologickými operátory	20
4.3.2	Segmentace metodou FCM.....	21
4.3.3	Segmentace distančním mapováním aktivní kontury	21
4.3.4	Segmentace použitím metody Gradient Vector Flow (GVF)	22
4.4	Analýza a klasifikace.....	23
4.4.1	Klasifikace použitím umělých neuronových sítí	23
4.4.2	Klasifikace pomocí Support Vector Machine.....	23
4.4.3	Barevná texturní analýza	24
4.4.4	Klasifikace jader metamyelocytů pomocí morfologických vzorů.....	25
5	Praktická část	26
5.1	Krevní nátěry	26
5.2	Pořizování snímků	26
5.3	Parametry struktur v obraze.....	27
5.3.1	Jádro leukocytů	27

5.3.2	Cytoplazma leukocytů	28
5.3.3	Velikost leukocytů	29
5.3.4	Erytrocyty	29
5.3.5	Trombocyty.....	30
5.3.6	Pozadí a artefakty.....	30
5.4	Programové segmentování buněčných oblastí	31
5.4.1	Barevné úpravy	32
5.4.2	Morfologické operace a filtrace	33
5.4.3	Segmentace jádra v obraze	33
5.4.4	Korekce oblasti jádra	34
5.4.5	Segmentace cytoplazmy	35
5.4.6	Chybovost segmentace jádra a cytoplazmy	39
5.5	Příznaky oblasti jádra a cytoplazmy	41
5.5.1	Analýza dat jádra a cytoplazmy.....	42
5.5.2	Statistický souhrn příznaků.....	43
5.5.3	Výběr příznaků	49
5.6	Programová klasifikace příznaků	50
5.6.1	Topologie sítě	50
5.6.2	Trénování sítě	53
5.6.3	Chybovost sítě.....	55
6	Rozvržení algoritmu v programovém prostředí MATLAB.....	57
6.1	Funkce main.m	57
6.2	Funkce jadro_seg.m a cytoplazma_seg.m	57
6.2.1	Funkce region.m	58
6.3	Funkce klasifikace.m a neuronové sítě.....	58
6.3.1	Neuronové sítě net1_b.mat, net2_b.mat a net3_b.mat.....	58
7	Diskuze	59
7.1	Barevné odchylky	59
7.2	Různé počty leukocytů	59
7.3	Důraz na výběr vzorových buněk.....	59
7.4	Možné rozšíření práce	60
8	Závěr	61
9	Slovník pojmů.....	63

10	Citovaná literatura	64
11	Doplňková literatura.....	66

Seznam příloh

Příloha 1	Charakteristiky Gaussových filtrů.....	68
Příloha 2	Testy normality intenzitních příznaků jádra a cytoplazmy	69
Příloha 3	Testy normality geometrických příznaků jádra a cytoplazmy	70
Příloha 4	Testy normality příznaků šikmosti a špičatosti jádra a cytoplazmy.....	71
Příloha 5	Korelační grafy příznaků šikmosti a špičatost histogramu cytoplazmy.....	72
Příloha 6	Korelační grafy intenzitních příznaků jádra a cytoplazmy	73
Příloha 7	Graf intenzitních příznaků cytoplazmy před korekcí	74
Příloha 8	Vybrané 2D grafy intenzitních příznaků cytoplazmy	75
Příloha 9	Případ výskytu více buněk v obraze.....	76

Seznam obrázků

Obr. 1	analyzátory Cellavision DM96 a DM1200, nátěrový automat Sysmex SP-1000i[2] ...	13
Obr. 2	myeloidní a lymfoidní vývojová linie[6]	14
Obr. 3	neutrofilní tyče	15
Obr. 4	neutrofilní segmenty	16
Obr. 5	eozinofilní segmenty	16
Obr. 6	bazofilní segmenty	17
Obr. 7	monocyty.....	17
Obr. 8	malé lymfocyty	17
Obr. 9	velké lymfocyty.....	18
Obr. 10	segmentace jádra (vlevo), segmentace cytoplazmy (vpravo) [12].....	20
Obr. 11	postup distančního mapování a GFV [9]	21
Obr. 12	sada vybraných typů buněk [15]	24
Obr. 13	vliv aditivního a impulzního šumu na přesnost klasifikace [14]	25
Obr. 14	postup zjednodušení obrazu, porovnání s morfologickými vzory[13]	25
Obr. 15	eozinofilní segment ve světlém poli (vlevo) a pod fázovým kontrastem (vpravo).....	26
Obr. 16	příklad výběru jádra monocyty a získané hodnoty z histogramu RGB	27
Obr. 17	trombocyty (a), anizocytóza trombocytů (b) a gigantické trombocyty (c, d)	30
Obr. 18	ukázka možných artefaktů	31
Obr. 19	blokové schéma segmentace jádra (nahore) a cytoplazmy (dole).....	31

Obr. 20 barevné prostory RGB (nahore) a HSV (dole), zleva neutrofily, lymfocyty monocyty eozinofily a bazofily	32
Obr. 21 po aplikaci <i>otevírání</i> (vlevo), filtrace Gaussovým filtrem (uprostřed), binární maska jádra (vpravo).....	34
Obr. 22 vstupní obrazy pro segmentaci cytoplazmy (viz postup výše) zleva: Y_1 , Y_2 , Y_3 , Y_4 .	36
Obr. 23 vypočtené obrazy Z_1 , Z_2 , Z_3 z předložených vstupů	36
Obr. 24 vypočtené obrazy ERY_1 , ERY_2 a jejich součet ERY	37
Obr. 25 obraz po zvýšení intenzity pixelů.....	37
Obr. 26 výsledky procesu zvýraznění cytoplazmy.....	38
Obr. 27 případ nežádoucí segmentace trombocytů v případě nepoužití funkce pro odstranění větších regionů	39
Obr. 28 příklad nedostatečně segmentovaných jaderných jader (tmavé ohraničené oblasti)	39
Obr. 29 příklad nadměrně segmentovaných jader (vnitřní ohraničené oblasti)	40
Obr. 30 příklad nadměrně segmentované cytoplazmy (světlé oblasti)	40
Obr. 31 příklad nadměrně segmentované cytoplazmy eozinofilů (oblasti mimo buňku)	41
Obr. 32 eozinofily s nerovnoměrným rozložením cytoplazmy	46
Obr. 33 neutrofily s toxickou granulací (vlevo fyziologický neutrofil)	46
Obr. 34 struktura navržené hlavní neuronové sítě (schématicky)	51
Obr. 35 transformační funkce, vlevo signum (skryté neurony), vpravo lineární transf. funkce (neurony 2. vrstvy).....	51
Obr. 36 příklad tvoření učebního vektoru 12 buněk z každé populace leukocytů pro hlavní neuronovou síť	54
Obr. 37 struktura celého algoritmu navrženého v MATLABu	57

Seznam tabulek

Tab. 1 procentuální zastoupení jednotlivých leukocytů v periferní krvi.....	18
Tab. 2 průměrné hodnoty intenzit RGB kanálů oblasti jádra.....	28
Tab. 3 průměrné hodnoty intenzit RGB kanálů oblasti cytoplazmy	28
Tab. 4 poměry cytoplazmy a jádra	29
Tab. 5 průměrné hodnoty intenzit RGB kanálů oblasti erytrocytů	30
Tab. 6 průměrné hodnoty intenzit RGB kanálů oblasti pozadí	30
Tab. 7 počty získaných buněk	42
Tab. 8 tabulka procentuálních chyb pro dva učební poměry (testovací množiny, 111 buněk)	55
Tab. 9 výsledky kompletní klasifikace třemi neuronovými sítěmi	56
Tab. 10 počet buněk chybně zařazených do jiných populací, šipky značí kam byly buňky nesprávně zařazeny	56

Seznam grafů

Graf 1 velikost oblastí jádra a cytoplazmy	29
Graf 2 statistický přehled průměrů intenzit barevných složek (R, G, B) jádra (N) a cytoplazmy (C)	43
Graf 3 intenzity barevných kanálů (R, G, B) oblasti jádra (N), vlevo: zdvojený shluk (barevná odch.), vpravo: po korekci	44
Graf 4 intenzita barevných kanálů (R, G, B) cytoplazmy (C).....	45
Graf 5 statistický přehled geometrických příznaků jádra (N) a cytoplazmy (C) v pixelech a jejich poměrů	47
Graf 6 zleva: obvod jádra (N) a cytoplazmy (C) pro každou buněčnou linii zvlášť	48
Graf 7 zleva: plocha jádra (N) a cytoplazmy (C) pro každou buněčnou linii zvlášť	48
Graf 8 vlevo: poměr obvodů cytoplazmy a jádra, vpravo: poměr ploch cytoplazmy a jádra (všechny populace)	48
Graf 9 rozdělení buněčných řad v prostoru na základě zvolených příznaků	52

1 Úvod

V dnešní době jsou na diagnostické laboratorní metody kladeny vysoké nároky z hlediska rychlosti analýzy, její senzitivity a specificity a také nároky finanční. V laboratorní hematologii je již dlouho vymezený trend, zaměřený na automatizované přístroje. Většina laboratorních metod je plně automatizována. Automatizace umožnila rozvoj a zkvalitnění laboratorní analýzy. Například průtokové cytometry, měří tisíc buněk během několika vteřin s vysokou přesností, nátěrové automaty dovedou na základě hematokritové hodnoty zhotovovat a barvit standardizované krevní nátěry pro další mikroskopické hodnocení. Automatizované koagulometry umí pomocí optických či mechanických principů stanovit poruchy srážecí kaskády.

Každá z těchto analýz vyžaduje ke své realizaci větší či menší množství reagensií (dle metody), se kterými vzrůstá závislost na výrobcí, dodržování expiračních lhůt, možné omezení analýzy atd. Výjimkou je zatím jedna z nejstarších kvantitativních metod, které postačí pouze mikroskop, imerzní olej, ručně i automaticky zhotovený krevní nátěr, jedná se o tzv. „diferenciální rozpočet leukocytů“. Tato metoda spočívá v subjektivním mikroskopickém rozlišení bílých krvinek dle morfologických znaků buněk. Zhodnocené buňky jsou počítány a řazeny do jednotlivých tříd (záleží na laboratoři a SOP, 5–7 tříd [1]). Tyto rozpočty jsou většinou prováděny na 100–200 buněk (procentuální zastoupení, relativní počet), nebo mohou být přepočteny na celkový počet bílých krvinek v jednom litru krve (WBC – absolutní počet, údaj z analyzátoru). Díky této metodě získá klinik potřebný údaj ke stanovení diagnózy a hodnocení průběhu onemocnění. Vzhledem k nenáročnosti a velkému klinickému významu metodu diferenciálního rozpočtu leukocytů aplikují i ty nejmenší laboratoře vybavené potřebnou technikou.

Vedle tohoto jednoduchého vyšetření se běžně provádí jeho rozšířené varianty, spočívající v zahrnutí širší škály hodnocených buněk červené i destičkové řady a jejich početních a morfologických abnormalit. Další specializace může být podpořena jiným druhem barvicí techniky (např. barvení na železo, barvení na průkaz specifických enzymů). K prokázání určitých onemocnění jsou hodnoceny nátěry kostní dřeně (nutný odborně kvalifikovaný pracovník). Morfologické vyšetřování nátěrů kostní dřeně je časově i odborně poměrně náročné díky nadměrné buněčnosti všech vývojových stádií nejen leukocytární řady, ale i linie červené a destičkové (megakaryocytární). Morfologické hodnocení kostní dřeně nelze hodnotit jako diferenciální rozpočet kvůli jeho povaze a odlišným zaměřením morfologie (viz slovník pojmů).

Metodu diferenciálního rozpočtu leukocytů se díky rozmachu digitálního zpracování obrazu podařilo automatizovat až v posledních několika letech. V České republice jsou ke strojové diagnostice používány převážně přístroje řady DM firmy Cellavision se sídlem Lund, Švédsko (Obr. 1)[2]. Ve stručnosti se jedná o analyzátory vybavené podavačem krevních

nátěrů a motorizovaným mikroskopem se záznamovým zařízením. Hardwarová stránka přístroje je precizně řešená avšak pro další zpracování málo podstatná. Hlavním stěžejním prvkem celé analýzy je výrobcem vyvinutý software s grafickým rozhraním a skrytým algoritmem pro hodnocení různých druhů fyziologických a patologických forem krevních elementů. Aby byla zachována standardizace sběru dat z různých vzorků, je vhodné používat strojově zhotovené krevní nátěry. Ručně zhotovené nátěry nemusejí vždy vyhovovat požadavkům přístroje na kvalitu provedeného mikroskopického preparátu. Navržené zpracování a hodnotící algoritmus je výsledkem mnohaleté práce softwarových designérů s předními morfologickými specialisty.[2; 3]



Obr. 1 analyzátory Cellavision DM96 a DM1200, nátěrový automat Sysmex SP-1000i[2]

Navrhnout kvalitní software, který obsáhne co největší škálu hodnocených elementů, není nikterak jednoduché. Požadavky směřují k vysoké senzitivitě a specificitě, samostatnosti celého systému při analýze a k intuitivnímu uživatelskému rozhraní umožňujícímu dodatečnou korekci výsledků klinikem či laborantem. Dalším podstatným bodem je schopnost rozlišení elementů, které nejsou za normálních okolností jednoznačně rozpoznatelné (překrytí, odchylky vzhledu, některé patologické stavy), kde musíme brát v potaz vysokou variabilitu vzhledu buněk u různých pacientů.

2 Cíl práce

Tato práce má za cíl shrnout a zhodnotit používané techniky a principy, které jsou nezbytné pro základní i pokročilé práce s obrazem k získání vhodného vstupního materiálu pro další zpracování a následnou diferenciaci krevních elementů. Po zmapování literárních pramenů následuje popis problematiky při zpracování obrazu leukocytů.

Prvořadým úkolem je nasbírat reprezentativní soubor pěti druhů krvinek z nátěrů periferní krve. V dalším kroku je nutné shromážděná data analyzovat a zaměřit se na separaci buněčných součástí (jádro a cytoplazma). Na základě segmentovaného jádra a cytoplazmy je možné extrahovat příznaky. Získané příznaky je nutné dále statisticky zpracovat a vybrat soubor příznaků vhodný pro následující klasifikaci.

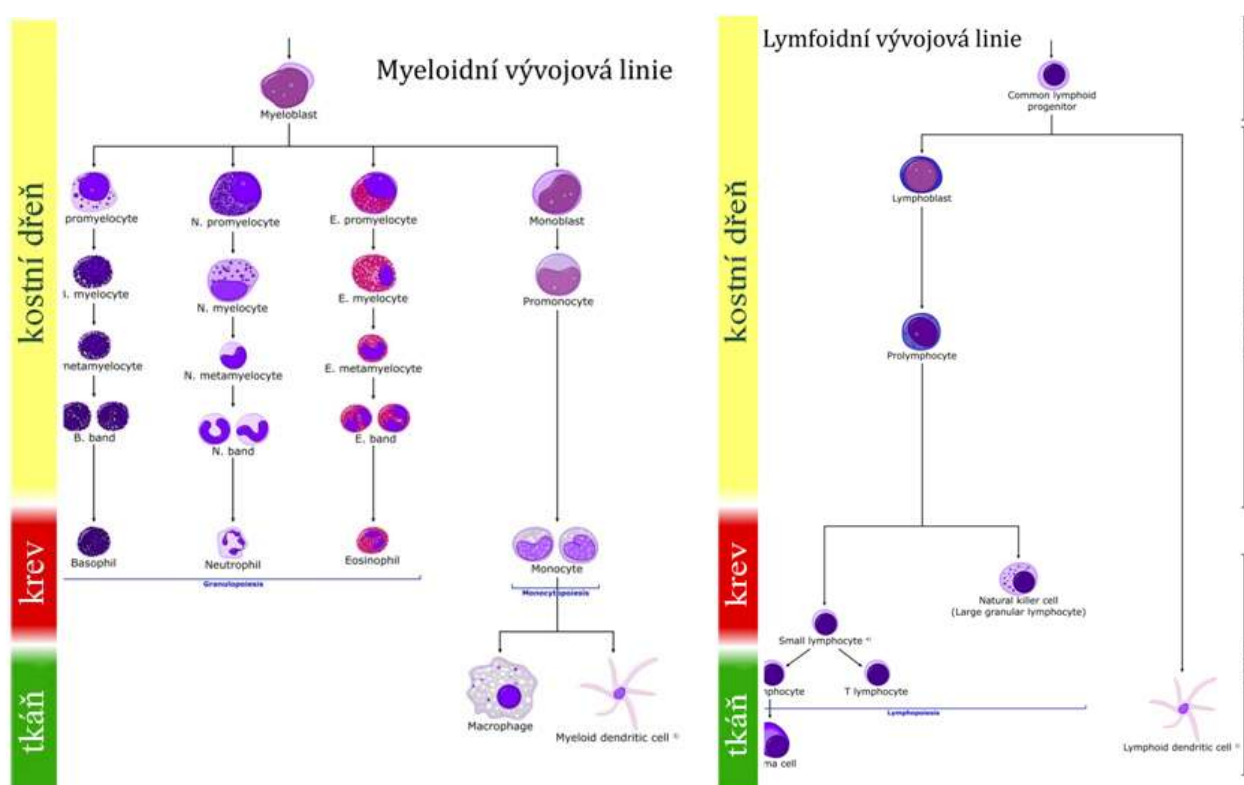
Jedním z dalších cílů je navržení klasifikačního procesu, který je schopen rozpoznat a korektně zařadit předkládané příznaky leukocytů. Jako klasifikační proces byla zvolena metoda klasifikace pomocí umělých neuronových sítí.

3 Hematologický základ

Pro navržení účinného postupu rozpoznávání buněk je v první řadě nutné se seznámit s morfologií leukocytů zastoupených v nátěrech periferní krve.

3.1 Charakteristika leukocytů

Leukocyty (bílé krvinky) jsou součástí buněčné složky krve. Uplatňují se zejména při obranyschopnosti organismu. Za fyziologických podmínek se nachází v lidské krvi v koncentraci $4-10 \times 10^9/l$ bílých krvinek. Leukocyty vznikají v krvetvorných orgánech. Granulocytární řada dává vznik efektorovým buňkám s obsahem specifických granul (Obr. 2). V monocytární řadě vznikají monocyty. Granulocytární a monocytární vývojová řada mají společnou prekurzorovou buňku a tím jsou zahrnuty v tzv. myeloidní vývojové linii (Obr. 2). Z prekurzorové lymfoidní buňky vznikají postupným vyžíváním v jednotlivých krvetvorných orgánech (kostní dřeň, thymus) všechny formy lymfocytů (lymfocyty T a B, NK buňky), (Obr. 2). Zralé nebo téměř zralé buňky opouští kostní dřeň, granulocyty a monocyty přecházejí do krevního oběhu, T a B lymfocyty do lymfatických systémů, kde dozrávají a jsou pak částečně vyplavovány do obvodové krve. V průběhu zrání se buňka leukocytu morfologicky mění. Cytoplazma během vývoje ztrácí svou bazofilii, u některých buněk (granulocyty) se objevují specifická granula. Jaderný chromatin se zahušťuje a ztrácí jemné uspořádání, u zralých granulocytů může segmentovat. Buňka postupně zmenšuje i svůj objem a mění tvar. [4; 5]



Obr. 2 myeloidní a lymfoidní vývojová linie[6]

3.2 Třídy leukocytů

K rozlišení bílých krvinek v nátěrech periferní krve využíváme některé jejich morfologické charakteristiky. Všímáme si zejména tvaru a vzhledu jádra, vyzrálosti jaderného chromatinu, vzhledu a zabarvení cytoplazmy, přítomnosti granul aj. Polymorfonukleáry (polynukleáry) mají laločnaté a členěné jádro. Řadí se k nim efektorové buňky granulocytární řady (viz slovník pojmů). Naopak mononukleáry mají na rozdíl od předešlé skupiny jádro jednoduší, kulovitého nebo oválného tvaru. Jiné dělení, které vychází z toho, zda cytoplazma obsahuje či neobsahuje granula, rozlišuje granulocyty a agranulocyty. Granulocyty od stádia zralého myelocytu obsahují specifická granula v cytoplazmě a v této řadě rozlišujeme neutrofilní, eozinofilní a bazofilní granulaci (viz slovník pojmů). Agranulocyty nemají specifickou granulaci, pouze u některých lymfocytů může být v cytoplazmě přítomno malé množství azurofilních granul. [4; 5]

3.3 Leukocyty hodnocené v barveném nátěru periferní krve

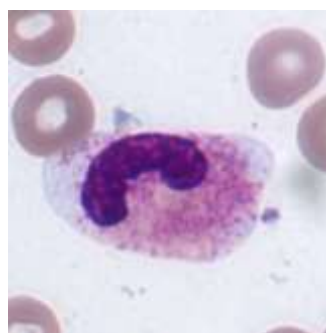
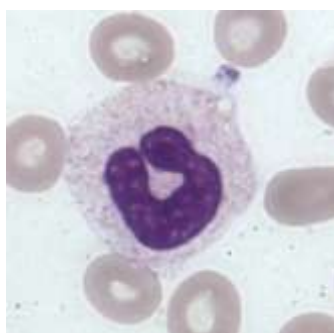
3.3.1 Nesegmentovaný granulocyt (tyč)

Průměr: 10–16 μm .

Jádro: není segmentováno, má tvar podkovy nebo kličky či rovné nebo lehce do oblouku stočené tyčky. Povrch jádra je pravidelný bez výraznějšího zaškrncení.

Chromatin: hutný, nerovnoměrně rozmístěn, má hrudkovitý charakter, tvoří tmavší a světlejší políčka.

Cytoplazma: většinou slabě oxyfilní s granulací.[7]



Obr. 3 neutrofilní tyče

3.3.2 Segmentovaný neutrofilní granulocyt (neutrofilní segment, neutrofil)

Průměr: 12–15 μm .

Jádro: zaškrncené většinou rozdělené na 2–5 částí (segmentů) spojených tenkými nitkovitými můstky.

Chromatin: hutný, tmavší políčka jsou seskupena po obvodu jádra.

Cytoplazma: oxyfilní s četným zastoupením specifických granul, která obsahují enzymy (neutrální proteázy, kyselé hydrolázy, myeloperoxidázu,

esterázy), cytotoxické a antimikrobní látky, cytoplazmatické membránové receptory.[7]



Obr. 4 neutrofilní segmenty

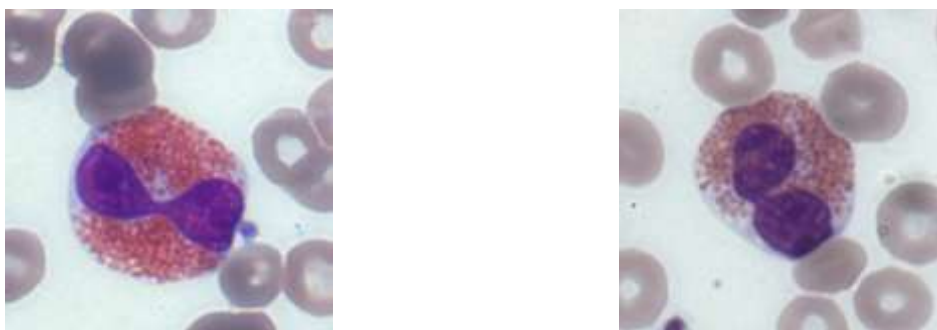
3.3.3 Segmentovaný eozinofilní granulocyt (eozinofilní segment, eozinofil)

Průměr: 10–16 μm .

Jádro: většinou segmentováno na 2 fragmenty propojené vzájemně chromatinovým můstkem. Vzhledem připomíná činku nebo brýle.

Chromatin: hutný, místy prosvětlený.

Cytoplazma: oxyfilně zabarvená s četným zastoupením specifických eozinofilních granul, která obsahují např. bazický protein, histaminázu, histamin, kyselé hydrolázy, eozinofilní peroxidázu, plazminogen.[7]



Obr. 5 eozinofilní segmenty

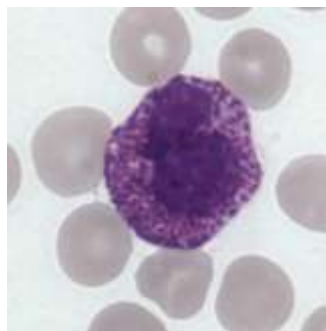
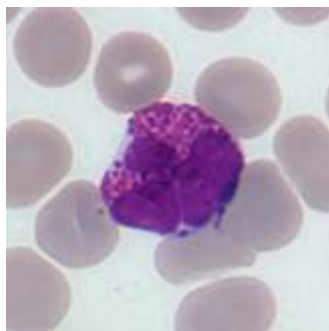
3.3.4 Segmentovaný bazofilní granulocyt (bazofilní segment, bazofil)

Průměr: přibližně 11 μm .

Jádro: laločnaté, připomíná jetelový list, často rozděleno na 3 segmenty, spojené mezi sebou silnějším chromatinovým můstkem. Jádro nemusí být mnohdy díky výrazné granulaci dobře viditelné.

Chromatin: obsažen v menší míře než u neutrofilního segmentu, tvoří hutná, tmavá políčka.

Cytoplazma: obsahuje specifická granula, hrubá a větší než u ostatních typů segmentů. Granula obsahují histamin, serotonin, prostaglandiny a heparin.[7]



Obr. 6 bazofilní segmenty

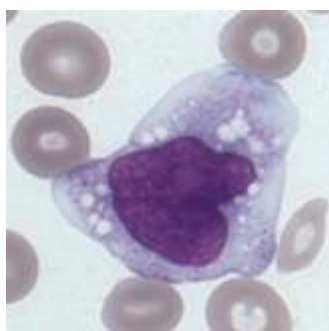
3.3.5 Monocyt

Průměr: přibližně stejný jako u promonocytu.

Jádro: nepravidelné, ledvinovité.

Chromatin: řídký, vláknitý s tmavšími políčky.

Cytoplazma: šedomodrá, často s vakuolami, s jemnou granulací, která není v běžném barvení příliš zřetelná.[7]



Obr. 7 monocity

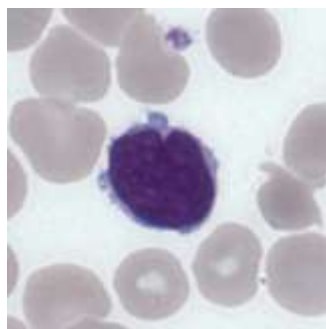
3.3.6 Malý lymfocyt

Průměr 8–10 μm .

Jádro: kulaté, vyplňuje velkou část buňky.

Chromatin: hutný, homogenní, jadérko není viditelné.

Cytoplazma: slabě až středně bazofilní, tvoří okolo jádra úzký lem.[7]



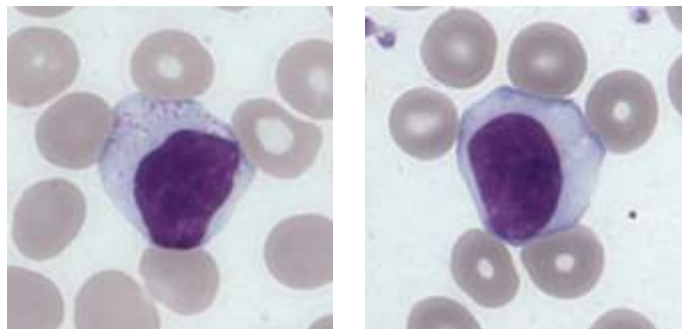
Obr. 8 malé lymfocyty

3.3.7 Velký lymfocyt

Průměr: 12–16 μm .

Chromatin: oproti malému lymfocytu řidší, mohou být i viditelná jadérka.

Cytoplazma: oxyfilní, nebo lehce bazofilní, v některých případech může obsahovat několik primárních azurofilních granul seskupených v jednom místě.
[7]



Obr. 9 velké lymfocyty

3.4 Procentuální zastoupení leukocytů v periferní krvi

Jelikož má každá laboratoř vlastní referenční rozmezí, referenční hodnoty leukocytů pocházejí z vnitřního standardu oddělení klinické hematologie II. Interní kliniky Fakultní nemocnice Hradec Králové a jsou zde uvedeny pouze jako doplňková informace.[7]

Buňka	Výskyt (dospělí)
Neutrofilní granulocyty	48–70 %
Lymfocyty	23–45 %
Monocyty	2–8 %
Eozinofilní granulocyty	1–4 %
Bazofilní granulocyty	0–2 %

Tab. 1 procentuální zastoupení jednotlivých leukocytů v periferní krvi

Tyto uvedené počty se mohou měnit v závislosti na lokálním obyvatelstvu a v závislosti na předpisech a standardech laboratoře. Tyto změny jsou pouze nepatrné.

4 Literární průzkum programového zpracování

Při porovnání všech pramenů se ukazují základní společné znaky použitých metod.

4.1 Využití barevných prostorů

Některé studie popisují využití různých barevných prostorů pro dosažení lepší segmentace a interpretace výstupu pro další analýzu obrazu. Ve většině případů jsou obrazy v RGB prostoru převedeny na stupně šedi a poté *binarizovány* [8; 9; 10; 11; 12; 13]. Zřídka se používá převodu do prostoru HSV či LAB [14]. Pokud je obraz ponechán barevný, může být vhodnější pro texturní zpracování či pro zpracování metodou SVM (Support Vector Machine). Frekventovanějším postupem je spíše jednotlivé barevné složky obrazu v LAB či HSV převést do šedotónové reprezentace. Tento převod může zvýraznit nebo potlačit kontrast potřebných struktur v obrazu či zjednodušit jejich interpretaci pro další manipulaci s obrazem.

4.2 Předzpracování

Pro snadnější práci se vstupními obrazy je vhodné je náležitě upravit. Fotografie vstupují do uvedených programových postupů ve většině případů ručně ořezané a upravené (jas, kontrast, gamma korekce). Pro zjednodušení jsou buňky oříznuty tak, aby na každém obrazu byla pouze jedna buňka. V některých člancích jsou zdrojové obrazy předloženy expertovi (často klinik, cytolog, hematologický specialista) pro korektní určení jejich typu. Tento krok slouží jako prostředek pro dodatečné zpřesnění výsledné diferenční analýzy. Obrazy jsou děleny na 2–3 struktury. Jsou jimi: jádro, cytoplazma, pozadí (ne vždy). Důraz je autory kladen na předzpracování snímků k získání vysoké kvality a jejich segmentaci. Pro separaci potřebných struktur jsou použity metody jako prahování obrazu s využitím znalosti o histogramu, distanční mapování, aktivní kontury, morfologické operátory, popřípadě jejich kombinace. Výsledky těchto technik slouží jako vstup pro klasifikační analýzu.

4.3 Segmentace

K dosažení segmentace samostatných leukocytů je nutné je co nejlépe separovat od ostatních elementů v obraze. K tomuto účelu slouží segmentační techniky, které jsou pevným základem pro další operace s obrazem. Segmentaci obrazu lze realizovat mnoha způsoby, často se používají lokální informace obrazu jako např. histogram, intenzita pixelu, shluky. Tyto techniky se dělí do dvou kategorií.

První kategorie se nazývá hledání regionů a je založená na prahování obrazu na základě histogramu šedotónového obrazu. Druhou kategorií je tzv. detekce kontur, která využívá nesouvislosti intenzity obrazu nebo textur u krajů objektu. Vzhledem k tomu, že ne vždy jsou okraje buněk dostatečně ostré, je obtížné vytěžit maximální informaci o hranách obrazu. Proto jsou často tyto techniky pro zvýšení efektivity kombinovány s matematickými

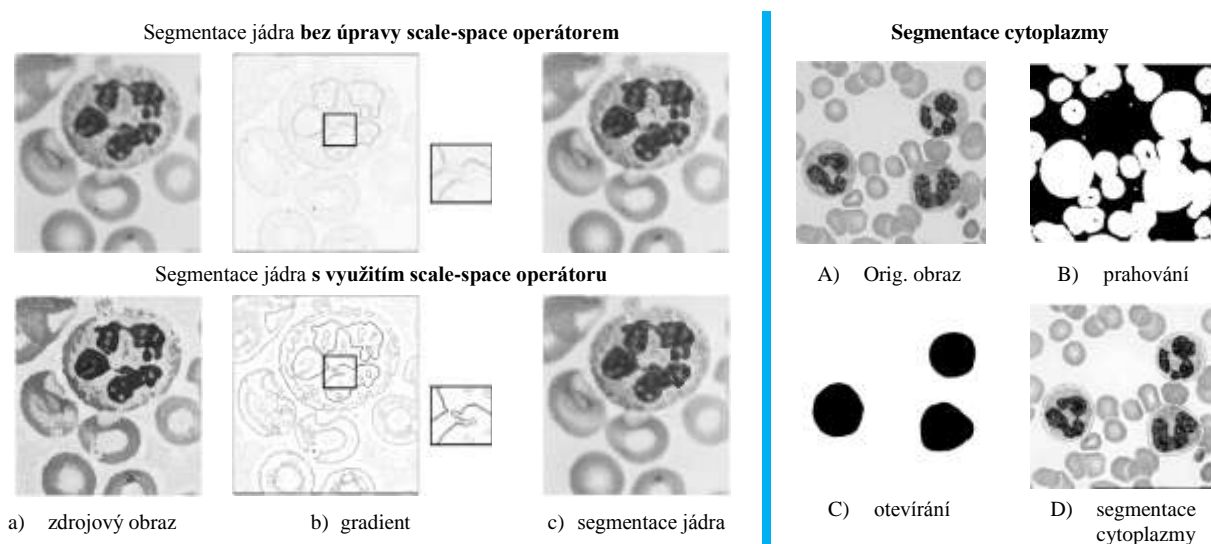
morfologickými operacemi, mezi nejfrekventovanějšími jsou eroze, dilatace, metoda *rozvodí* (watershed).

4.3.1 Segmentace matematickými morfologickými operátory

Jedna ze studií metod segmentace leukocytů využívá k segmentaci jádra následující postup. Začátek algoritmu zpracovává obraz základními operacemi (úprava jasu a kontrastu). Dalším stupněm je prahování, které je důležité pro pozdější fáze zpracování obrazu. Následně je originální obraz „zjednodušen“ tzv. *scale-space* operátorem (Laplacián operátor, prostorově závislá dilatace a eroze, Obr. 10a). Vzniká zjednodušený obraz, který sebou nese původní geometrickou-tvarovou podstatu objektů, avšak tímto omezením je zbavený jemných texturních detailů, které jsou pro další zpracování nepodstatné. [12]

Díky neostrosti a nepravidelnosti okrajů jádra je následně na zjednodušený obraz aplikována eroze obrazu. Tímto krokem je dosaženo odstranění malých částí cytoplazmy těsně naléhajících na jádro, nebo regionů pixelů umístěných mezi jeho segmenty. Pro výpočet gradientu a ohraničení jádra buněk zjednodušeného obrazu je využita metoda Watershed, která je vedena prahovaným obrazem (Obr. 10b, c). [12]

K oddělení cytoplazmy od pozadí využívají autoři článku morfologické operace (granulometrická funkce) zaměřené na eliminaci erytrocytů (dle jejich velikosti a šíře distribuce v obraze). Pozadí, které je zpravidla nejsvětlejší oblast obrazu, je od všech buněčných elementů separováno prahováním (Obr. 10B). Následuje použití eroze, která pohltí erytrocyty na základě jejich velikosti oproti leukocytům. Posledním krokem je pak dilatace (charakterizovaná jako *otevírání*, (Obr. 10C)), která zvýrazní cytoplazmu buněk. [12]



Obr. 10 segmentace jádra (vlevo), segmentace cytoplazmy (vpravo) [12]

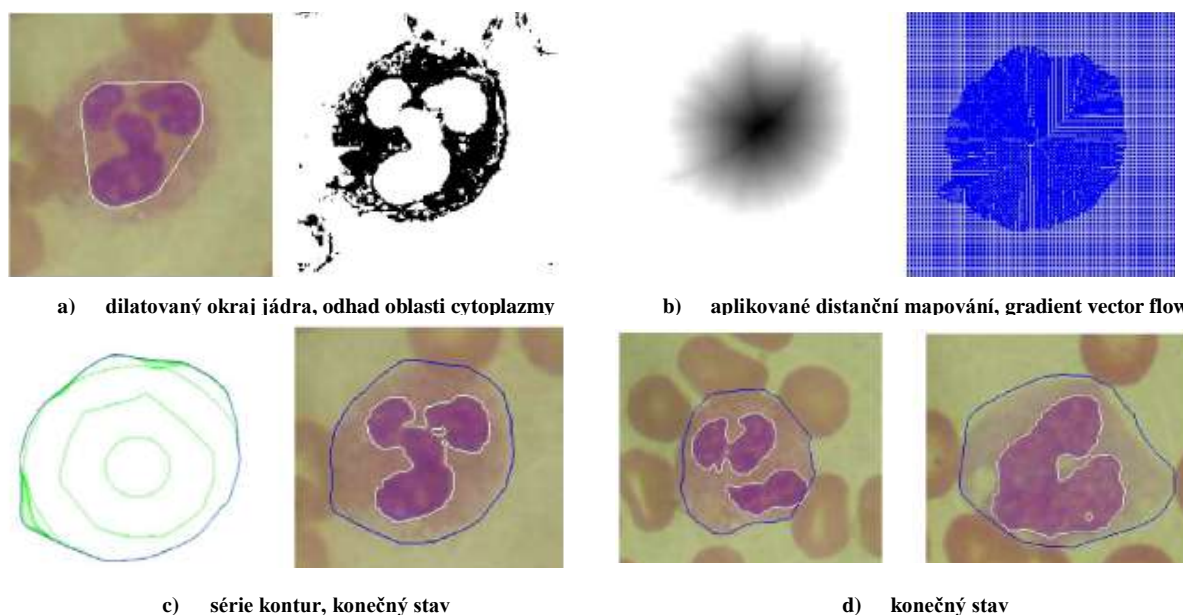
4.3.2 Segmentace metodou FCM

Segmentační technika Theera Umpona využívá automatický postup pro segmentaci cytoplazmy a jádra užitím metody fuzzy C-means clustering (dále FCM) a základních morfologických operátorů. Šedotónové snímky jsou kvůli nekonzistenci buněčných komponent filtrovány mediánovým filtrem. Na tento obraz je poté aplikována FCM metoda, která ho rozdělí na jednotlivá centra shluků. Podle zvolených kritérií (např. pokud je centrum shluků menší než průměr všech tmavých center, jedná se o jádro) rozdělí obraz na regiony jádra, cytoplazmy a pozadí. Region pozadí je pro další zpracování nepotřebný a je z FCM vyřazen.[10; 11]

Na regiony jádra a cytoplazmy je aplikována morfologická operace uzavírání. Jako konečná fáze segmentace je označována vzájemná kombinace obou oblastí. Pro zpřesnění úspěšnosti segmentační techniky byla použita pravděpodobnost chyby. Automaticky segmentované buňky byly porovnány s buňkami ručně segmentovanými odborníkem, tímto krokem byla provedena korekce vah FCM analýzy. [10; 11]

4.3.3 Segmentace distančním mapováním aktivní kontury

Tato použitá technika se zaměřuje na nalezení konvexních okrajů jádra. Barevný obraz je *binarizován*, pak prahován a oříznut tak, aby zachytil celou buňku. Následně je potřeba odhadnout okraje jádra. Zde je aplikována tzv. convex hull (konvexní slupka-obálka) technika, která se zdá vhodná pro odhad jader leukocytů s komplikovaným tvarem jádra (neutrofily, eozinofily a bazofily). *Konvexní obálka* jádra je autory článku definována jako nejmenší konvexní množina bodů ve tvaru polygonu, které obsahují či obklopují jádro. Pomocí techniky *Hit-or-Miss* je nalezen konvexní tvar jádra. [9]



Obr. 11 postup distančního mapování a GVF [9]

Díky tomu, že jádro zralého leukocyty je obklopeno cytoplazmou, je využita obálka jádra, která je dilatována na základě skupiny pixelů uvnitř dilatovaného prstence (Obr. 11a). Tato skupina pixelů pak tvoří referenční tabulku pro odhad cytoplazmy. Primární oblast cytoplazmy je tak hrubě odhadnuta pro další zpracování (Obr. 11a).[9]

Stěžejním postupem konečné segmentace jádra a cytoplazmy je metoda aktivních kontur, která zdůrazní oblast jádra a cytoplazmy. Obraz primární cytoplazmy je změněn pomocí matematických morfologických operací (vyplňování, uzavírání). Takto upravený obraz (binární forma) je pak transformován na distanční mapu (Obr. 11b). Dále je odvozen tok gradientu (Obr. 11b) a později využit pro zlepšení dynamiky aktivní kontury (Obr. 11c).[9]

Eozinofily a lymfocyty, na kterých jsou erytrocyty napojeny na povrch cytoplazmy, jsou touto metodou bez problému správně ohraničeny. Monocyty a neutrofily nemají často správně vyznačenou cytoplazmu. Aktivní kontura tak vyhodnocuje nesprávně místa s erytrocyty blízko cytoplazmy, nebo v případě *konturování* v okolí blízko okraje cytoplazmy (Obr. 11d). Bazofilní granulocyty jsou vzhledem k vysoké granularitě hůře rozlišovány v oblasti jádra a cytoplazmy. Dle autorů by tyto chyby mohly být eliminovány přesnějším odhadem primární cytoplazmy. Upravení barvicího procesu by mělo zlepšit barevnou nekonzistenci v okolí cytoplazmy leukocytů. [9]

4.3.4 Segmentace použitím metody Gradient Vector Flow (GVF)

Autoři studie nejprve převedou originální obraz do šedotónové reprezentace. Poté provedou ruční separaci leukocytů jejich výřezem. Takto předpřipravený obraz poslouží jako vstup pro rozostření Gaussovým filtrem. V obraze tak dojde k redukci šumu a potlačení nechtěných detailů. Dále je vypočten gradient pomocí operátoru (např. Robersova, Sobelova, Prewittové). Gradientní obraz je prahován tak, aby všechny prvky hran byly zachovány a většina šumu byla potlačena. [8]

Použitý postup kontroluje, zda je každý nenulový pixel větší než jeho dva sousední ve směru gradientu. Pokud pixel podmínce nevyhovuje, je nulován. V opačném případě je ponechán. Pomocí dvou prahovacích kritérií jsou rozeznány *slabé* a *silné* hrany. Silné hrany jsou ponechány a spojeny. V poslední fázi segmentace jádra je použita samotná GVF metoda, která využívá externí síly (gradient či drobná modifikace) řídící konturu. Dochází tak k obklopení jádra leukocyty. Takto získané křivky poslouží jako informace pro matematickou morfologickou operaci označovanou jako *vyplnění*. [8]

Pro segmentaci cytoplazmy poslouží výřez leukocyty ze začátku předchozí segmentace, ze kterého je pomocí subtrakce odečtena oblast jádra (výstupní obraz segmentace jádra). V obraze jsou tak zastoupeny složky cytoplazmy, erytrocytů a pozadí. Nyní je nutné oblasti jádra a erytrocytů od sebe oddělit. Z hlediska jasové intenzity těchto dvou oblastí je

vhodné použít segmentační techniku. V této studii byl konkrétně použit Zackův algoritmus (*prahování histogramu*). [8]

Autoři uvádějí průměrnou úspěšnost segmentace jádra 92 % a 78 % pro segmentaci cytoplazmy oproti ručně segmentovaným leukocytům. Algoritmus byl aplikován pro diagnostiku onemocnění, kde je patrná odlišná morfologie buněčných řad (leukemie).[8]

4.4 Analýza a klasifikace

Většina navržených technik následuje manuální proces detekce, extrakce vlastností a klasifikace buněk. Za pilotní krok, který je společný pro většinu klasifikačních metod, se považuje extrakce příznaků. V tomto bodě je vhodné do analýzy dodat předzpracované či zjednodušené obrazy (jasové úpravy, filtrace, segmentace, binarizace, viz předchozí kapitola). Příznaky jsou získávány přímo z naměřených parametrů (geometrické a barevné vlastnosti: plocha, obvod, průměr, barva), nebo jejich vzájemným odvozením (statistické parametry: průměrná odchylka, parametry histogramu, šikmost, špičatost). Zřídka se objevuje texturní analýza, která je omezena vzhledem k variabilitě, mnohočetnosti buněk a rozlišení snímků.

4.4.1 Klasifikace použitím umělých neuronových sítí

Beksac a kol. dosáhli 61% míry klasifikace použitím umělé neuronové sítě a prostorově-barevně závislých příznaků 16 tříd buněk v kostní dřeni a periferní krvi. Ongun a kol. použili barevné snímky 12 typů leukocytů kostní dřeni a periferní krve. Pro dosažení dobré senzitivity a specificity využili 57 příznaků z jádra a cytoplazmy založených na intenzitě, barvě, textuře a tvaru. Bylo využito 70 % buněk pro učení a 30 % pro testování neuronové sítě, kdy docházelo k náhodnému výběru z každé třídy. Celý proces byl 100krát zopakován, aby bylo získáno nejlepší rozhodovací kritérium. [11]

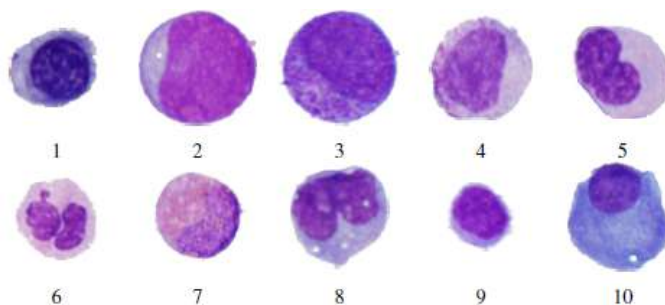
Využitím vícevrstvého perceptronu bylo dosaženo 90% míry správné klasifikace. Výběrem pouze nejlepších výsledků z testovací skupiny dochází k celkovému nadhodnocení jejich metody. Tyto skutečnosti vnášejí určité pochybnosti o možném zobecnění této metody. Použitím stejného podílu buněk, sloužících jako vstup (učební a zkušební skupiny buněk) jsou klasifikátory ovlivněny tím, že některé buněčné složky se normálně vyskytují v odlišném poměru vzhledem k ostatním.[11]

4.4.2 Klasifikace pomocí Support Vector Machine

Touto metodou detekují autoři článku 10 tříd buněk (Obr. 12). Po segmentaci obrazu bylo nutné se zaměřit na možné měřitelné parametry buněk. Data byla nashromážděna ze souboru čítajícího více než 6 000 vzorků. Bylo vybráno 7 parametrů zahrnujících: oblast jádra, poměr plochy buňky a jádra, symetrii, kompaktnost jádra a vlastnosti histogramu červeného kanálu buňky (průměr, variabilita, a šikmost). Těmito parametry byly vymezeny prahy, do kterých náleží pouze vybrané třídy buněk. Dalším krokem je generování

diagnostických parametrů, charakterizujících jednotlivé buněčné typy mezi sebou. Tyto definované vlastnosti tvoří tři skupiny (texturní, geometrickou a statistickou). [15]

Texturní příznaky jsou popisovány jako změna intenzity stupňů šedé v prostoru. Pro popis texturních parametrů byly zvoleny *Unserovy parametry a Markovova náhodná pole deskriptorů*. Geometrické parametry korespondují s proporcemi jádra a celé buňky. Jedná se o obvod, poloměr, jejich poměr, plochu, kompaktnost, konkavitu, počet konkávních bodů atd. Statistické parametry byly získány z charakteristik histogramů RGB obrazu. Bylo vydefinováno více jak 150 parametrů. [15]



Obr. 12 sada vybraných typů buněk [15]

Vzhledem k velkému počtu parametrů, možné nejednoznačnosti a vlivu šumu na některé z nich, je hlavním problémem jejich výběr. Autoři vybírají faktory na základě eliminace multidimenzionality, kdy zůstanou právě nejsilnější faktory, které spolu nekorelují. Dosažení maximálních výsledků je důležité pro další učení klasifikace systému. V tomto ohledu se nejlépe osvědčila metoda linear Support Vector Machine. [15]

4.4.3 Barevná texturní analýza

V této studii A. I. Plastinin a kol. využívají statistické vlastnosti lokálních texturních charakteristik. Vzhledem k vzájemné podobnosti barevných kanálů RGB (autory popisováno jako křížová korelace) je nutné použít tzv. nekorelovaný barevný prostor (Rudemanův *laβ* prostor). *Laβ* prostor vychází z analýzy komponent LMS prostoru. [14]

Překlad RGB prostoru do *laβ* je realizován ve dvou krocích, kdy se prostor RGB maticovým součinem s konstantním vektorem převede do LMS a ten následně obdobným způsobem do *laβ*. Pro zajištění statistických charakteristik získaného prostoru jsou výsledky normalizovány do intervalu $\langle 0;1 \rangle$, přičemž je každý kanál normalizován dle průměru a standardní odchylky.[14]

Vyvinutá statistická metoda je založena na 3-dimenzionální distribuční funkci pravděpodobnosti v obraze. Porovnáváním skupiny bodů ve všech kanálech jsou hledány hodnoty, které vstupují do distribuční funkce. Získanými číselnými parametry pro všechny 3 barevné složky obrazu jsou korelace, stíny, kontrast a inertnost. Tyto proměnné jsou následně vstupem pro diskriminační analýzu, která najde nejlépe vystihující faktory a potlačí

přebytečné parametry. Na základě diskriminační analýzy byly použity příznaky, které maximalizují zvolená kritéria pro oddělení jednotlivých tříd leukocytů. [14]

Vybrané faktory jsou od sebe lineárně oddělitelné. Proto autoři uvažují možnost vytvoření lineárního klasifikátoru po částech na základě metody SVM. V konečné fázi je diskutováno o přesnosti klasifikace s ohledem na rušení obrazu aditivním nebo impulzním šumem (Obr. 13). [14]

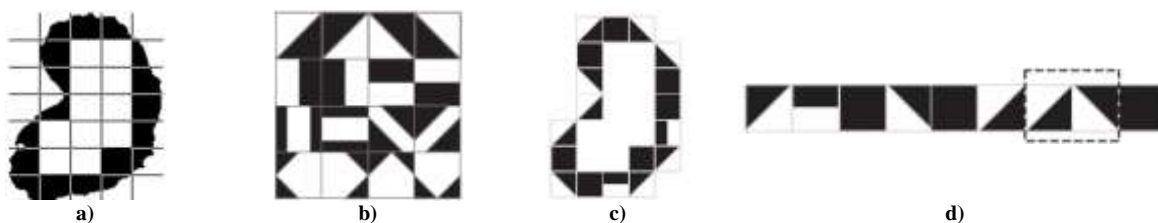


Obr. 13 vliv aditivního a impulzního šumu na přesnost klasifikace [14]

4.4.4 Klasifikace jader metamyelocytů pomocí morfologických vzorů

Tento případ klasifikace je do této práce zahrnut pouze okrajově. Jedná se o klasifikaci jádra metamyelocytů nezralých buněk granulocytární linie (metamyelocytů). Autoři vycházejí z charakteristických rysů jádra buňky, tím je specifický ledvinovitý tvar a zářez (zaškrcení) na jedné straně jádra. Do algoritmu vstupují binární obrazy již segmentovaných jader. Pro další manipulaci s obrazem je důležitá hraniční reprezentace, která je rozdělena do 3 tříd: parametrické kontury, hraniční body, aproximační křivky (geometrická primitiva reprezentovaná rovnými čarami aproximujícími povrch jádra). Pomocí aproximace úsečkami je možné označit charakteristické sekvence průběhu křivek jádra. [13]

Jedním z možných řešení přístupu je získání lineární reprezentace obrysu po částech a vyhodnocení strmosti v jednotlivých úsecích. Segmentovaný obraz je rozdělen na několik čtvercových podobrazů (každý si ponechává své 3 složky hraniční reprezentace, Obr. 14a). Jsou využity pouze podobrazy korespondující s povrchem. Každý jednotlivý prvek je porovnán s předdefinovanými morfologickými vzory (Obr. 14b). Vzniká tak zjednodušený obraz reprezentovaný úsečkami (Obr. 14c). V dalším kroku nastává extrakce příznaků. Jsou porovnávány vždy sousedící podobrazy a hledá se specifická část charakterizující zářez na jádře (Obr. 14d). Autoři aplikovali metodu v programovém prostředí MATLAB a udávají 100% úspěšnost identifikace na předložených metamyelocytech, udávají však malý počet testovaných obrazů.[13]



Obr. 14 postup zjednodušení obrazu, porovnání s morfologickými vzory[13]

5 Praktická část

5.1 Krevní nátěry

Krevní nátěry jsou použity z předcházející bakalářské práce. Jedná se o fyziologické nátěry periferní krve zhotovené OKH Fakultní nemocnice Brno Bohunice. Nátěry byly zhotoveny a obarveny automatem Sysmex SP-1000i (sídlo firmy: Kobe, Japonsko). Použitím strojového zhotovení a barvení je dosažena uniformita každého nátěru. Uniformita je zajištěna přesným nastavením natíracích podmínek pro měnící se parametry krve. Před natíracím procesem přijme automat od průtokového cytometru naměřené informace o krevním vzorku související s hustotou a viskozitou krve (hematokrit a počet erytrocytů). Dle získaných hodnot je nastaven objem dávkované krve na podložní sklíčko. Dále je nastaven sklon roztěrového skla a délka natírací dráhy. K barvení byla použita technika Giemsa Romanowski – May Grümwald (viz slovník pojmů). Jedná se o základní hematologické barvení. Barvicí proces je přesně časově rozvržen, aby nedocházelo k případným barevným odchylkám. I přes zajištění jednotnosti v barvě nátěrů může docházet k barevným odchylkám, této problematice se věnuje kapitola 5.3.6.

5.2 Pořizování snímků

Snímky byly pořizeny kamerou Nikon DS-Fi1 připojenou na mikroskop stejné značky (Eclipse 80i). Stonásobného zvětšení bylo dosaženo použitím imerzního objektivu. Snímky mají formát TIFF o velikosti 2560×1920 pixelů. Prostorové rozlišení bylo změřeno kalibračním sklem o škále 0,01 mm a odpovídá přibližně 135 pixelům na 0,01 mm neboli 13,5 $\mu\text{m}/\text{pix}$. Před snímáním bylo provedeno nastavení celé apertury. Zdroj světla byl zvolen na předdefinovaný režim pro fotografování. Clona zdroje světla byla otevřená a clona kondenzoru byla uzavřena na poloviční průsvit. Pořizování snímků probíhalo v programu Nikon NIS Elements ver. 2.2. Jelikož byl mikroskop vybaven několika typy zobrazovacích módů (světlé pole, tmavé pole, fázový kontrast), bylo možné pořizovat fotografie více způsoby (zvoleny: světlé pole a fázový kontrast, Obr. 15).



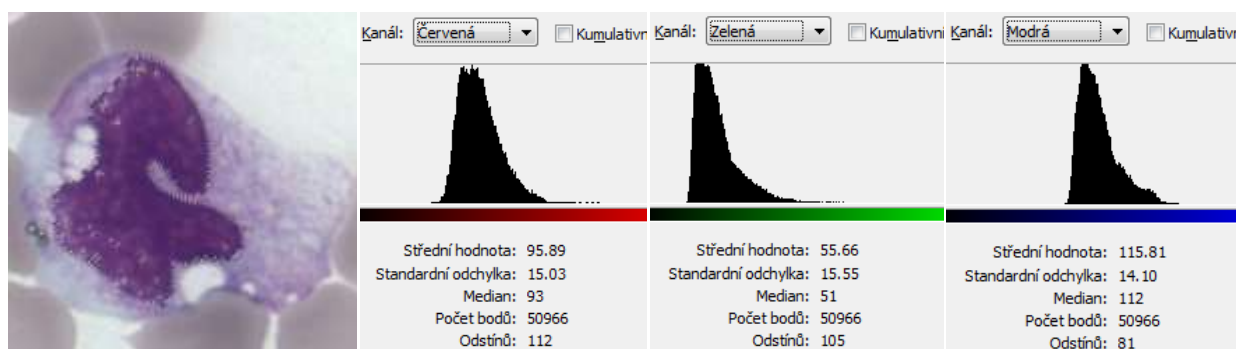
Obr. 15 eosinofilní segment ve světlém poli (vlevo) a pod fázovým kontrastem (vpravo)

Při hledání vhodné expozice dosahovaly hodnoty příliš krátkých časů, mezi kterými se expoziční doba volila obtížně. Tento problém byl odstraněn použitím filtru ND8. Nastavení expozice bylo 300 ms pro světlé pole a 30 ms pro fázový kontrast. Gamma korekce byla nastavena na hodnotu 1,4krát a položka kontrast nastavena na lineární (ostatní módy nabýzely

nelineární nebo neodhadnutelné nastavení kontrastu). Při fotografování bylo dosaženo maximální ostrosti v obraze díky sledování parametru maximálního kontrastu ve scéně. Fotografování pod fázovým kontrastem bylo prováděno pouze otočením kondenzorové karusele, tím bylo zajištěno lícování obrazů z obou módů.

5.3 Parametry struktur v obraze

Pro dostatečně kvalitní separování leukocytů v obraze je potřeba zjistit několik parametrů, které je odlišují od ostatních obrazových struktur. V následujících odstavcích jsou uvedeny tabulky základních statistik. Statistiku byly prováděny v grafickém editoru Zoner Photo studio (verze 12 Trial) ručním výběrem oblastí buněk a jednotlivých prvků obrazu. Pro každý měřený region byly vyhodnoceny parametry velikosti oblasti v pixelech a histogramy pro 3 barevné kanály včetně popisné statistiky (Obr. 16). Hodnoty byly zapisovány do tabulky v MS Excel (verze 2007 Standard) a následně sjednoceny průměrováním do určených tabulek. Vzhledem k orientačnímu účelu bylo takto zmapováno pouze 5 buněk z každé buněčné linie. Dalšími statistickými parametry se zabývá kapitola detailněji věnovaná extrakci příznaků.



Obr. 16 příklad výběru jádra monocyty a získané hodnoty z histogramu RGB

5.3.1 Jádro leukocytů

Jádra leukocytů se jeví jako nejtmavší struktura v obraze. Jejich tvar je různý pro každý typ. **Lymfocyty** (zkr. LYMF) Malé nebo Velké) mají jádro kulaté a pravidelné, menších rozměrů s kompaktním chromatinem. **Monocyty** (zkr. MONO) mají vzhledem k ostatním vyzrálým formám největší a nepravidelné jádro tvaru oválu, ledviny, listu apod. Chromatin u monocyty je nehomogenní, vláknitý s četnými prosvětlenými místy. Vzhledem k povaze chromatinu se může stát, že okraje jádra nejsou jasně ohraničené jako u ostatních forem a místy přecházejí plynule do cytoplazmy. **Neutrofily** (zkr. NEU) obsahují nepravidelné jádro rozdělené na 3–5 segmentů, jádro má hutný chromatin a nejsou v něm světlá místa. **Eozinofily** (zkr. EO) mají jádro většinou rozdělené na dva pravidelné segmenty, jeho barva je díky prosvítající cytoplazmě mírně zabarvena do červena. **Bazofily** (zkr. BAZO) mají jádro laločnaté, sytě fialové, často překryté velkým množstvím bazofilních granul. Často nejsou obrysy jádra dostatečně patrné.

Vhodnými parametry pro analýzu se zdají být plocha, obvod, průměr, těžiště, lokální zakřivení povrchu, barevnost. V Tab. 2 je uveden statistický souhrn variability barvy u jednotlivých buněčných tříd. Pro názornou ukázkou jsou buňky v tabulce obarveny výslednou barvou z kanálů RGB. Z hodnot je patrné, že nejsvětlejší jádro mají monocyty a nejtmaší malé lymfocyty. Eozinofily mají barevný odstín jádra mírně červený (výskyt specifické granule).

Parametr	Medián			Standardní odchylka		
BARVA	R	G	B	R	G	B
NEU	83	45	112	17,2	13,5	12,6
LYMFO M	70	36	99	9,1	9,5	10,8
LYMFO V	105	36	125	13,0	14,3	12,9
MONO	112	48	135	17,3	17,1	14,9
EO	96	48	121	14,8	12,2	11,1
BAZO	80	37	105	15,4	10,5	11,2

Tab. 2 průměrné hodnoty intenzit RGB kanálů oblasti jádra

5.3.2 Cytoplazma leukocytů

Cytoplazma leukocytů je odlišná od všech struktur v obraze. U granulocytů je barva cytoplazmy světlá, neutrofilní (viz slovník pojmů). Při pohledu jsou přítomné tmavé odstíny z důvodu zastoupení granul specifických pro daný typ.

Neutrofilní granula mají drobnější strukturu než následující, jsou sytě modrá a četná, zpravidla pokrývají celou plochu cytoplazmy. **Eozinofilní granula** mají střední velikost a jsou typicky zbarvená oranžovými až cihlově červenými odstíny, nemusejí pokrývat celou plochu buňky (občas tvoří shluky). **Bazofilní granula** jsou největší ze tří jmenovaných, dosahují nejtmaších odstínů, často masivně překrývají jádro, které lze hůře rozpoznat.

Cytoplazma velkých lymfocytů je světle modrá, občas jsou přítomny granula (v rámci jednotek). Malý lymfocyt má cytoplazmu pouze jako úzký lem okolo jádra, její barva dosahuje až sytě modrých odstínů. Monocyty jsou buňky s největším množstvím cytoplazmy, která je světlá, šedomodrá s výraznou nehomogenitou, často v ní mohou být bílé vakuoly či slabá granule.

Parametr	Medián			Standardní odchylka		
BARVA	R	G	B	R	G	B
NEU	184	156	186	19,6	27,3	17,4
LYMFO M	117	116	170	20,1	22,7	16,8
LYMFO V	160	155	200	17,0	21,6	13,7
MONO	180	165	210	19,8	25,8	14,6
EO	161	112	138	22,0	37,6	31,7
BAZO	130	80	140	33,5	33,0	20,9

Tab. 3 průměrné hodnoty intenzit RGB kanálů oblasti cytoplazmy

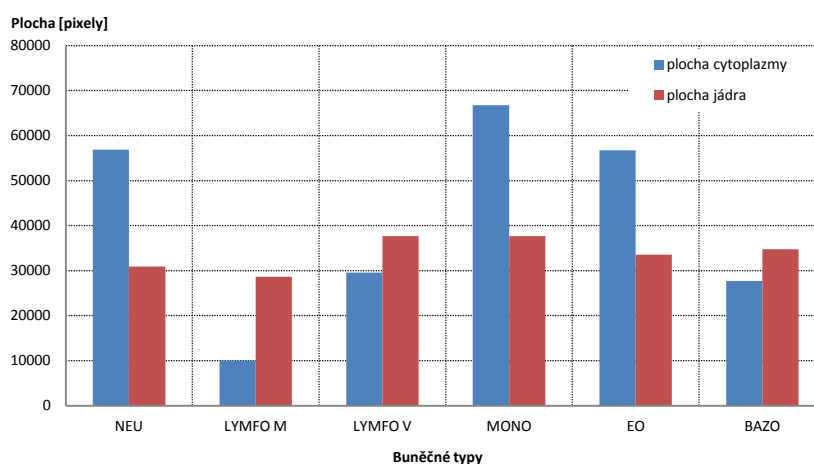
Tab. 3 popisuje barevnost cytoplazmy. Nejtmavší odstíny má cytoplazma bazofilů a malých lymfocytů. Nejsvětějšími odstíny jsou zbarveny cytoplazmy velkých lymfocytů a monocytů. Eozinofily a bazofily mají oproti ostatním buněčným typům poměrně velké standardní odchylky zbarvení jádra, což by mohlo způsobovat potíže při hledání prahu pro segmentaci. *Jelikož jsou rozdíly v barvě cytoplazmy viditelně významné, jeví se tato vlastnost buněk jako vhodný příznak pro klasifikaci.*

5.3.3 Velikost leukocytů

Graf 1 ukazuje, že leukocyty dosahují převážně větších rozměrů než normocyty (erytrocyt o fyziologické šířce cca 8 μm). Výjimkou se mohou stát malé lymfocyty a některé bazofily. Největšími buňkami mezi zmiňovanými 5 typy jsou monocyty. Tab. 4 popisuje poměry mezi jádrem a cytoplazmou v obraze u jednotlivých leukocytů. Neutrofil, eozinofil a monocyt mají přibližně stejný poměr cytoplazmy a jádra. Velké lymfocyty a bazofily mají rovněž stejný poměr. Malé lymfocyty patří mezi nejmenší leukocyty a jejich cytoplazma je v obraze zastoupena nejméně.

	Cytoplazma	Jádro
NEU	65%	35%
LYMFO M	26%	74%
LYMFO V	44%	56%
MONO	64%	36%
EO	63%	37%
BAZO	44%	56%

Tab. 4 poměry cytoplazmy a jádra



Graf 1 velikost oblastí jádra a cytoplazmy

5.3.4 Erytrocyty

Erytrocyty jsou bezjaderné buňky s hnědočervenou cytoplazmou. Fyziologické erytrocyty mají kulatý tvar a bikonkávní profil, což se v obraze projevuje prosvětlením středu buňky. Výjimkami mohou být starší nebo patologicky změněné buňky (sférocyty – menší erytrocyty tvaru koule). Vzácně se objevující patologické erytrocyty nejsou brány jako porucha. Jako patologické stavy jsou brány nátěry obsahující nad 5 % nebo nad 20 % defektních erytrocytů (závislé na patologickém stavu). Tab. 5 popisuje barevnost cytoplazmy erytrocytů, která se hodnotami jasu podobá jasové úrovni cytoplazmy leukocytů.

Existuje mnoho bizarních tvarů červených krvinek, např. kapkovité, tvaru ježka či pľlměsíce atd. Průměr erytrocytů se pohybuje okolo 8 μm , přičemž mohou dosahovat jak

miniaturních rozměrů ($< 6,5 \mu\text{m}$ – mikrocyty), tak i větších proporcí ($> 8,5 \mu\text{m}$ – makrocyty). Pokud krevní nátěr obsahuje směs různě velkých erytrocytů, jedná se o anizocytózu.

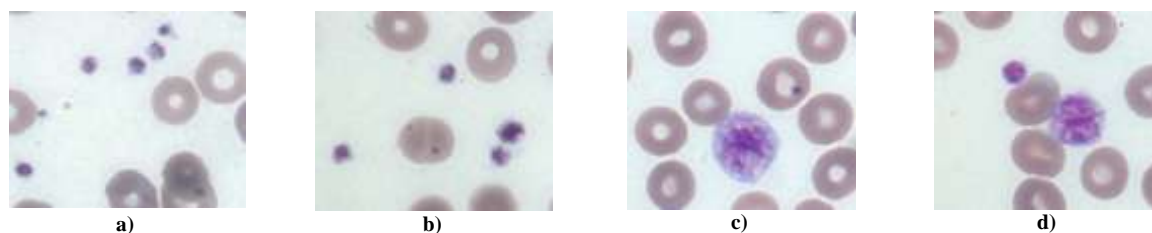
Parametr	R	G	B
Medián	162	171	157
Standardní odchylka	12,12	11,3	12,8

Tab. 5 průměrné hodnoty intenzit RGB kanálů oblasti erytrocytů

5.3.5 Trombocyty

V obraze se jeví jako tmavě modré roztřepené kuličky s projasněnými místy. Nemají jádro, jsou to ve skutečnosti fragmenty cytoplazmy megakaryocytů (vývojová forma). Barevností se blíží barvě jader leukocytů. Ve fyziologickém stavu mají průměr $2\text{--}4 \mu\text{m}$ a nepřesahují velikost erytrocytů (Obr. 17a, b).

V nátěru se mohou vyskytovat i tzv. obří trombocyty (Obr. 17c) nebo směs různých velkých trombocytů (vše nefyziologické stavy, Obr. 17d). Při poruchách srážení se můžou v nátěrech objevovat shluky trombocytů.



Obr. 17 trombocyty (a), anizocytóza trombocytů (b) a gigantické trombocyty (c, d)

5.3.6 Pozadí a artefakty

Tab. 6 dokazuje, že pozadí je nejsvětlejší struktura v obraze. Obklopuje všechny buňky. Pozadí přejímá barvu použitého světelného zdroje v mikroskopu. Tato barva je často pozměněna odstínem barviva, navázaného na ulpělé bílkoviny krevní plazmy mezi buňkami.

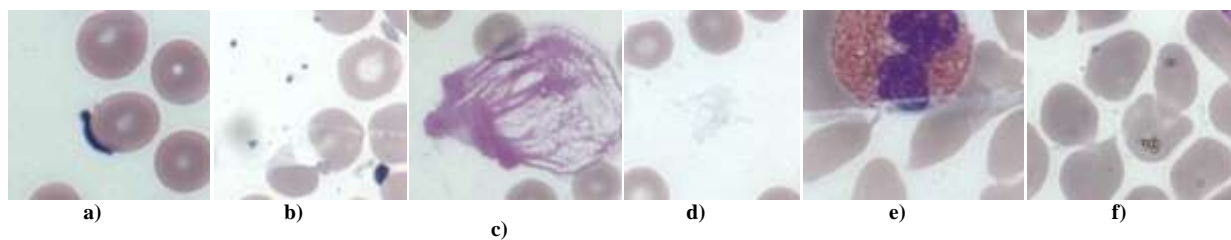
Parametr	R	G	B
Medián	221	229	229
Standardní odchylka	8,8	9,1	10,4

Tab. 6 průměrné hodnoty intenzit RGB kanálů oblasti pozadí

Mezi artefakty lze řadit necelé, rozpadlé buňky a jejich části. Pokud jsou barvicí chemikálie dlouho používané, mohou se v nátěrech objevovat různé fragmenty vysráženého barviva (Obr. 18a, b). Barvivo mohou ve zvýšené míře navázat některé druhy bílkovin krevní plazmy (Obr. 18c, d).

V případě časté manipulace s nátěrem vznikají v nátěru rýhy (Obr. 18e). Jestliže je nátěr plný drobných bublinek (Obr. 18f), došlo pravděpodobně ke špatné fixaci natřené krve, nebo se do fixačního roztoku (metanol) mohla přimísit voda. Další artefakty v obraze vznikají

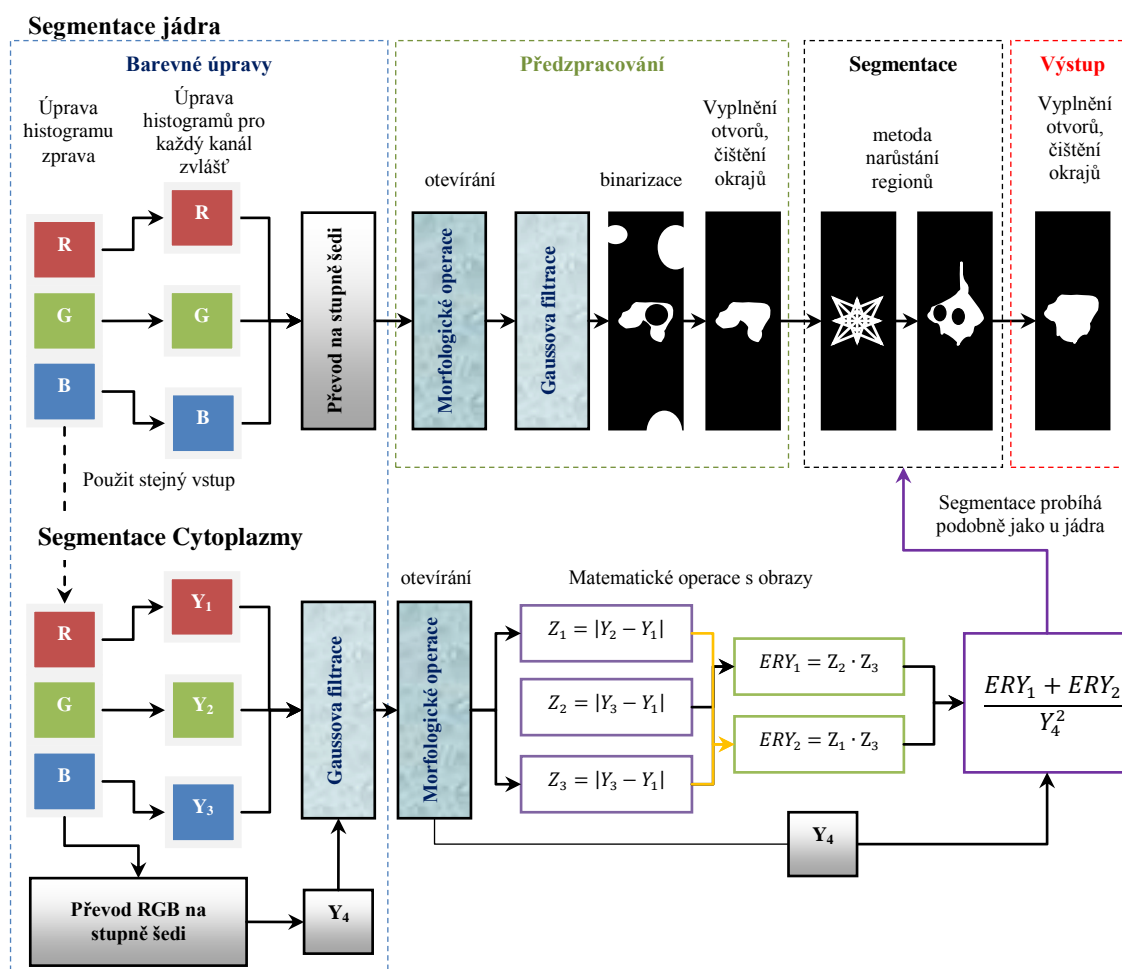
znečištěním optické dráhy v mikroskopu. Prach se projevuje jako tmavá zrna různých velikostí, mastné skvrny snižují lokální ostrost obrazu.



Obr. 18 ukázka možných artefaktů

5.4 Programové segmentování buněčných oblastí

Nasnímané fotografie je nutné pro klasifikaci vhodně zpracovat. K tomuto účelu slouží segmentace cytoplazmy a jádra, které jsou stěžejní pro získání příznaků. Pro zrychlení výpočtů a testování byly použity předem ořezané fotografie separátních leukocytů (600x600 pixelů), ovšem bez změny prostorového rozlišení. Blokové schéma na Obr. 19 popisuje celý proces segmentace jádra a cytoplazmy. Jelikož jsou některé bloky u segmentace cytoplazmy podobné segmentaci jádra, jsou u bloku segmentace cytoplazmy sloučeny podobné procesy.

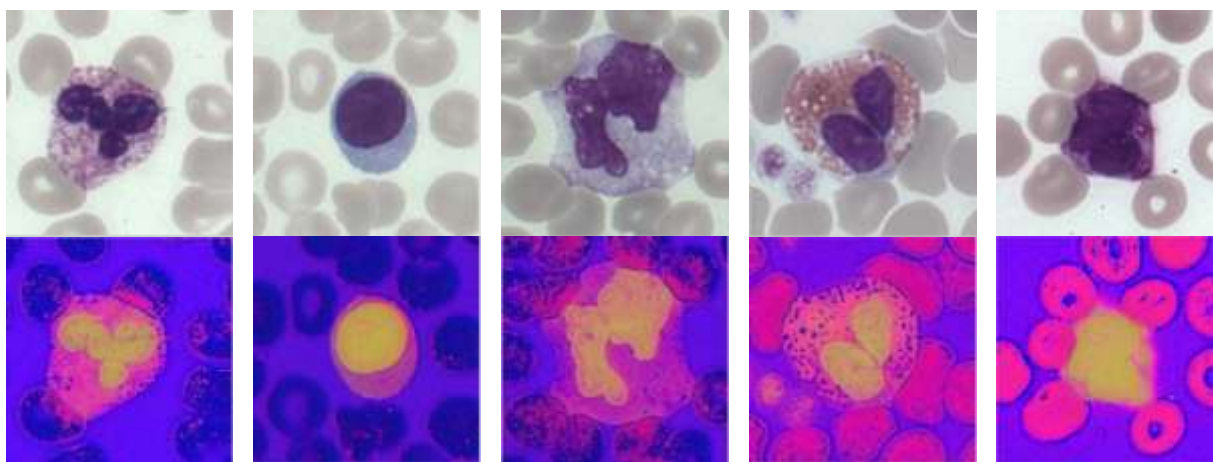


Obr. 19 blokové schéma segmentace jádra (nahore) a cytoplazmy (dole)

5.4.1 Barevné úpravy

Výše uvedené statistické zpracování potvrdilo, že jádro je zpravidla nejtmavším elementem obrazu. Pro dobré výsledky je potřeba jádro od okolí separovat. Byly vyzkoušeny prahovací techniky v prostorech RGB, LAB a HSV (Obr. 20) pro jednotlivé barevné složky. Zdrojová data byla zaznamenána v RGB prostoru, do zbylých dvou byl nutný převod pomocí přepočtu. U LAB prostoru byl přepočet nejnáročnější a trval déle než u HSV. Prostory LAB a HSV měly výrazný problém v reprezentaci jádra a cytoplazmy.

U některých populací leukocytů (monocyty a eozinofily) nebylo jádro nejtmavším objektem a tudíž bylo problematické stanovit podmínky pro účinnou separaci jádra od okolí. Jelikož je HSV odvozen z RGB prostoru jen přepočtem, jsou si oba prostory do jisté míry podobné, avšak u HSV jsou odlišně barevně reprezentované.



Obr. 20 barevné prostory RGB (nahore) a HSV (dole), zleva neutrofil, lymfocyt, monocyty, eozinofil a bazofil

Jelikož byly fotografie pořízeny manuálně, mohlo dojít k drobným odlišnostem v opětovném nastavení mikroskopu, proto byla do postupu úprav zahrnuta úprava barevného rozsahu. Jedná se o oříznutí histogramu zprava z původního rozsahu $\langle 0; 255 \rangle$ na rozsah $\langle 0; 248 \rangle$ jednotně pro všechny barevné kanály. Touto úpravou je docílena normalizace osvětlení pozadí, které eliminuje drobné odchylky v jasu obrazu.

Omezení histogramu zleva nebylo provedeno, jelikož by došlo k ztmavení tmavých partií buňky (tmavé cytoplazmy a jádra) a poté by mohlo vzniknout zkreslení při klasifikačním procesu. Možnost adaptivity úpravy jasu histogramu automatizovanou funkcí na základě šířky histogramu v tomto případě nebyla zvolena z důvodu nutnosti zachování jednotných podmínek pro konečnou klasifikaci.

Nejefektivnější realizací se osvědčila úprava intenzity obrázku pro každý barevný kanál zvlášť (obdobně jako v předchozí korekci jasu). Princip spočívá ve výběru stanoveného rozsahu jasových hodnot v obrázku a přepočítání tohoto rozsahu na nově zvolený rozsah (zde

ponechán interval $\langle 0; 1 \rangle$ v datovém typu `double` resp. $\langle 0; 255 \rangle$ v celočíselné reprezentaci). Tato operace zvýší kontrast tmavších struktur v obraze (leukocyty, erytrocyty).

Postup:

1. Po načtení obrazu do pracovního prostoru MATLABu probíhá zpracování pomocí funkce `imadjust` jako normalizovaná úprava histogramu (viz výše) na všech barevných složkách stejně.
2. Pak následuje obdobná úprava jasu (fce `imadjust`) za účelem zvýraznění tmavého jádra separátně a specificky pro každý barevný kanál.
3. Pro zesílení efektu je tento postup jednou zopakován. Nastavení intervalů bylo nutné stanovit napevno (odhadem) a nastavit tak, aby vyhovovalo pro všechny zastoupené třídy buněk.
4. Jako poslední krok dochází ke sjednocení barevných kanálů do jednoho šedotónového dle vztahu (1), (fce `rgb2gray`).

$$\text{šedotónový obraz z RGB} = 0,2989R + 0,5870G + 0,1140B \quad (1)$$

5.4.2 Morfologické operace a filtrace

Pro zajištění korektního průběhu rozpoznání jádra bylo nutné snímky po kontrastní úpravě dále předzpracovat. Z důvodu dosažení lepších výsledků musel být získaný obraz zjednodušen a zbaven přebytečných detailů, které by mohly komplikovat další postupy. Zejména u granulárních forem bylo těchto manipulací potřeba. Morfologická operace *otevírání* funguje jako *eroze* následovaná *dilatací* (Obr. 21)[11; 12]. Strukturním elementem byl zvolen „disk“ o průměru 5 pixelů (0,4 μm). Průměr byl stanoven na základě velikosti granul a drobných elementů v cytoplazmě. Takto zvolená morfologická operace byla aplikována na šedotónový obraz k eliminaci struktury v obraze menší než její velikost. Pokud by byl použit disk o větším průměru, mělo by to za následek nežádoucí eliminaci některých částí jádra (můstky, prosvětlená místa, zakřivení). Morfologická operace *otevírání* byla také použita pro eliminaci drobných oblastí v binárním obraze jádra i cytoplazmy (viz kapitola 5.4.5), nebo k odstranění oblastí naléhajících na okraj obrazu. Některá jádra leukocytů mohou v binárním obraze jádra obsahovat otvory (nehomogenita nebo bizarní tvary, obsah vakuol). Eliminace otvorů v obraze probíhá pomocí morfologické operace *vyplňování* (filling).[8]

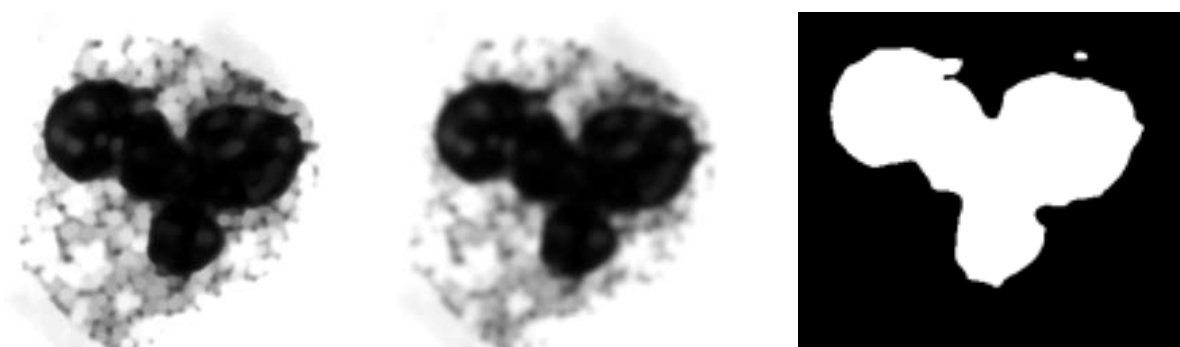
5.4.3 Segmentace jádra v obraze

Po aplikaci morfologických operací je třeba rozostřit drobné hrany vzniklé po kruhovém operátoru. K tomuto účelu slouží Gaussův filtr o velikosti 9 pixelů (0,67 μm)

a parametru sigma 5 (strmost Gaussovy křivky, přílohy). Filtr vhodně rozostří ostré přechody a obraz je tak připraven pro nalezení binární masky jádra (Obr. 21).

Postup:

5. Na obraz je aplikována morfologická operace *otevírání* (fce *imopen*), (Obr. 21).
6. Funkcí *imfilter* je provedena filtrace Gaussovým filtrem (viz výše).
7. Dále je nalezena minimální intenzita v obraze a záměrně nadhodnocena přičtením konstanty 0,25; aby byly zahrnuty pixely o vyšší intenzitě než jen minimum. Nalezením pixelů vyhovujících této podmínce je vytvořena binární maska, která odpovídá oblasti jádra.
8. Pokud jsou v bílé masce přítomny otvory, jsou vyplněny morfologickou operací *vyplňování* (fce *imfill*). Aplikací morfologické operace *otevírání* na binární masku (fce *bwareaopen* a *imclearborder*) jsou odstraněny drobné oblasti o celkové ploše 4 000 pixelů. Z obsahu kružnice bylo odvozeno, že kružnice o ploše 4 000 pixelů má průměr přibližně 72 pixelů a po přepočtu na délkovou jednotku přibližně 5,3 μm . Tato plocha byla stanovena podle velikosti fragmentů, nejčastěji trombocytů, shluků barviva. Tyto tmavší objekty většinou nesouvisí s jádrem (zbytky granulace, tmavší partie cytoplazmy), nebo naléhají na okraj obrázku.



Obr. 21 po aplikaci *otevírání* (vlevo), filtrace Gaussovým filtrem (uprostřed), binární maska jádra (vpravo)

5.4.4 Korekce oblasti jádra

Získaný binární obraz jádra je vzhledem k málo přizpůsobivému prahování pouze odhadem, proto je třeba jej zkorigovat. Některé buňky mohou mít i po kontrastní úpravě nejasné, mnohdy světlejší nebo rozostřené okraje jádra. Korekce spočívá ve využití dosavadní masky a expanze pomocí metody *narůstání regionů*, kde binární maska slouží jako počátek. Maska poté expanduje v jednotlivých cyklech do prostoru.

Prostorové šíření je realizováno osmi směry. V každém cyklu dojde k expanzi o jeden pixel ve všech směrech a pomocí rozhodovacího algoritmu je buď zahrnut do dosavadní masky (hodnota 1), nebo vyloučen (hodnota 0). Na konci každého cyklu je v binárním obraze

provedena *detekce hran Cannyho operátorem*, která slouží pro monitorování velikosti obvodu. Pokud se obvod od předchozího kroku nezmění, je korekce segmentace jádra ukončena. Rozhodování je uskutečněno porovnáním jasové hodnoty daného pixelu v obraze s adaptivním prahem.

Postup:

9. využívání *osmibodového okolí* je vhodné pro rychlou expanzi masky do prostoru, kde je porovnáván každý pixel s prahovou hodnotou. Do cyklu vstupuje hrubý odhad prahu dle určeného vztahem (2), kde \tilde{x} vyjadřuje medián a σ standardní odchylku kontrastně upraveného obrazu X a pod binární maskou M hodnoty 1.

$$THR_{j\acute{a}dro} = \min(X(M \in 1)) + 0.2 + \tilde{x}(X(M \in 1)) + \sigma(X(M \in 1)) \quad (2)$$

Poslední fází je vyplnění otvorů vzniklých po korekci jádra pomocí *otevírání*. Separované jádro je takto připraveno pro extrakci příznaků potřebných pro konečnou klasifikaci jednotlivých druhů buněk.

5.4.5 Segmentace cytoplazmy

Zvýraznění cytoplazmy

Segmentace cytoplazmy je rámcově podobná předchozí segmentaci jádra, avšak s některými změnami souvisejícími se zvýrazněním cytoplazmy. Cytoplazma je problematická část obrazu. V histogramu je umístěná různě, v závislosti na buněčné linii. Nejtmavší typy cytoplazmy (lymfocyty, bazofily) jsou dobře separovatelné, jelikož jsou na levém kraji histogramu a jasově nezasahují do jiných struktur obrazu (erytrocyty, pozadí). Cytoplazmy umístěné blíže ke středu histogramu (eozinofily) jsou problematické, protože mohou zasahovat na pozici histogramu, kde se nachází jasové složky erytrocytů. Monocyty a neutrofilní granulocyty mají mnohdy cytoplazmu světlou, která v histogramu těsně sousedí s pozadím. Z těchto důvodů je obtížné najít jednotnou separační funkci pro všechny buňky najednou.

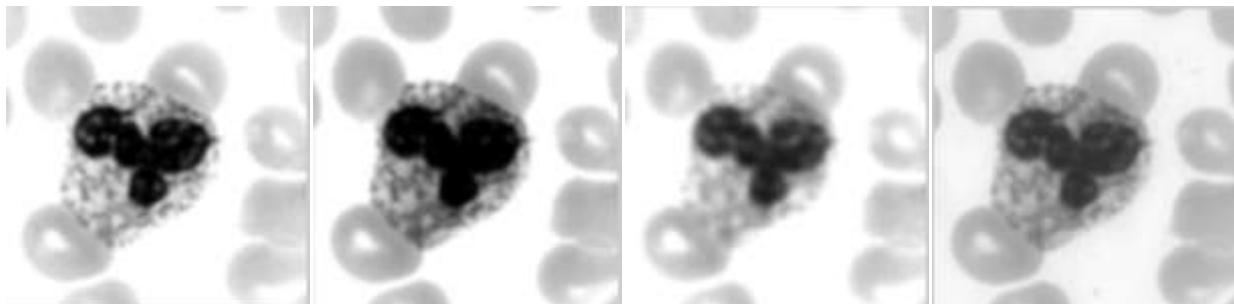
Mezi počáteční úpravy patří stejně jako u segmentace jádra jasová úprava realizovaná opět úpravou rozsahů původního barevného obrazu. Vzhledem k velké barevné rozmanitosti cytoplazmy zkoumaných buněk nejsou tyto kontrastní úpravy oproti segmentaci jádra tak razantní a mají za úkol výrazně zesvětlit pozadí a ztmavit cytoplazmu. Proto je tato úprava provedena pouze jednou.

Postup:

1. Do algoritmu vstupují všechny 3 složky *RGB* originálního obrazu (Y_1 , Y_2 , Y_3) a je u každé zvlášť provedena jasová úprava (fce *imadjust*). Dále do algoritmu vstupuje

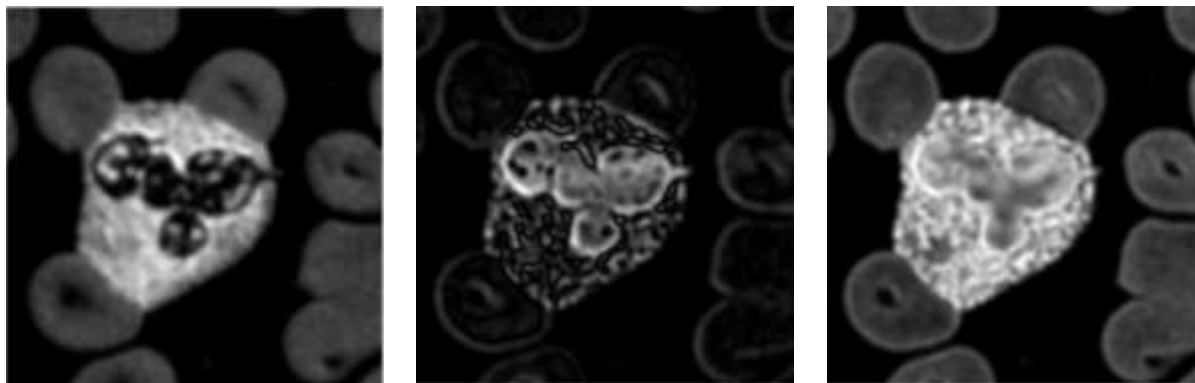
původní obraz (Y_4), jehož 3 barevné složky byly sloučeny do jednoho obrazu ve stupních šedi (fce `rgb2gray`). Jasové úpravy se u tohoto obrazu neprovádějí.

2. Filtrace Gaussovým filtrem o velikosti 17 pixelů ($1,3 \mu\text{m}$) a parametru sigma 7 (Příloha 1) má za úkol eliminovat všechny detaily a zachovat jen jednoduchý vzhled cytoplazmy (fce `imfilter`, viz Příloha 1). Filtrace je aplikována na všechny obrazy (Y_1, Y_2, Y_3, Y_4), (Obr. 22).



Obr. 22 vstupní obrazy pro segmentaci cytoplazmy (viz postup výše) zleva: Y_1, Y_2, Y_3, Y_4

3. V další fázi jsou vypočítány rozdíly absolutních hodnot obrazů dle vztahu 3 (Obr. 23). Tyto nově vzniklé obrazy jsou připraveny pro následující zvýraznění cytoplazmy a potlačení erytrocytů, které znemožňují korektní segmentaci cytoplazmy. Celkem by bylo možné tímto způsobem vypočítat 9 obrazů. Uvedené tři obrazy v nadcházejícím výpočtu podávají nejlepší výsledky. Obr. 23 ukazuje, že Z_1 a Z_3 jsou si podobné až na odlišnou barvu jádra a Z_3 má více obsažených okrajových kontur erytrocytů. V obrazu Z_2 je zvýšený zástup kontur, zbytek buňky kromě jádra a granulace je potlačen.



$$Z_1 = |Y_2 - Y_1|$$

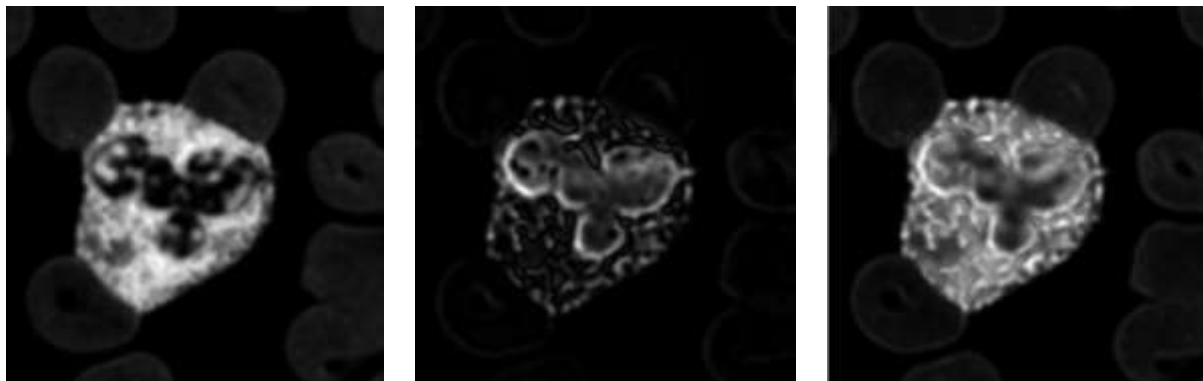
$$Z_2 = |Y_3 - Y_1|$$

$$Z_3 = |Y_3 - Y_2| \quad (3)$$

Obr. 23 vypočtené obrazy Z_1, Z_2, Z_3 z předložených vstupů

4. Vypočtené rozdíly obrazů jsou takto připraveny pro další operaci a to součin obrazů dle vztahu 4. Vynásobením obrazů je docíleno toho, že erytrocyty jsou potlačeny a v obraze zůstanou pouze části leukocytů.

5. Dále se provede součet obrazů ERY_1 a ERY_2 , tím je docílena vzájemná kombinace těchto dvou obrazů. V tomto případě je zesvětleno jádro a tmavší oblasti cytoplazmy. Sčítání se provádí také z důvodu sjednocení, protože eozinofily a lymfocyty mají odlišné barevné složení cytoplazmy.



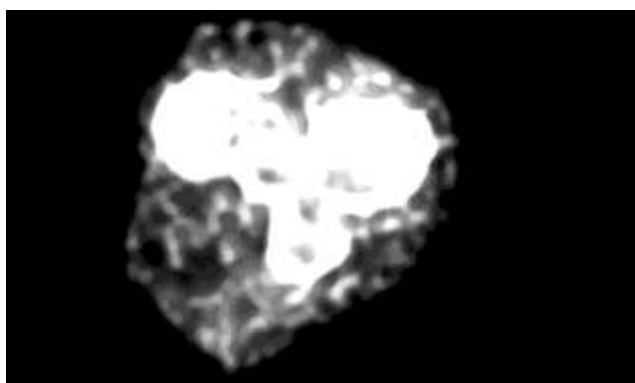
$$ERY_1 = Z_2 \cdot Z_3$$

$$ERY_2 = Z_1 \cdot Z_3$$

$$ERY = ERY_1 + ERY_2 \quad (4)$$

Obr. 24 výpočtené obrazy ERY_1 , ERY_2 a jejich součet ERY

6. Jelikož při součinu obrazů (pixely jsou v hodnotách double) dochází k silnému snížení jasových hodnot pixelů, je třeba je jistým způsobem zesílit. K tomuto účelu se dle vztahu 5 dosavadní obraz ERY podělí kvadrátem původního obrazu Y_4 (stupně šedi).
7. Kvůli výskytu ostrých přechodů mezi odstíny cytoplazmy (hlavně u granulárních forem leukocytů) je na upravený obraz ERY aplikována morfologická operace otevírání s maskou o průměru 3 pixely ($0,2 \mu\text{m}$).



$$ERY = \frac{ERY}{(Y_4)^2} \quad (5)$$

Obr. 25 obraz po zvýšení intenzity pixelů

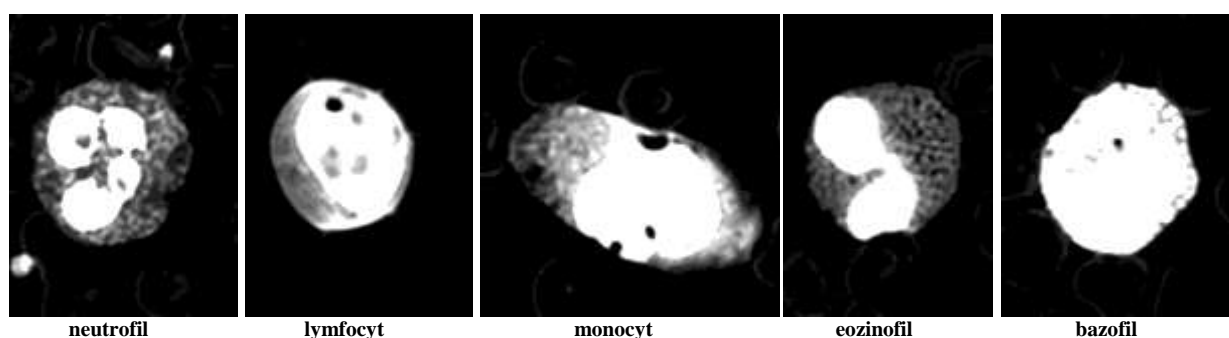
Výhodou této operace je potlačení okolních erytrocytů a zvýraznění cytoplazmy. Je zvýrazněno také jádro, avšak není na něho brán zřetel. Cytoplazma je tak zvýrazněna a její ohraničení záleží na dalším zpracování, zejména pak na nastavení prahových hodnot.

Za **nedostatek** může být považováno, že je třeba vkládat fotografie nafocené jedním způsobem. Tím je myšleno konstantní nastavení během každého snímání. V opačném případě mohou být výsledky horší.

Výsledky zvýraznění cytoplazmy

Výsledky zvýrazňujícího procesu ukázaly, že aplikované operace se příznivě projeví u všech tříd leukocytů. Nejlepší výsledky zvýraznění cytoplazmy podaly lymfocyty a bazofily.

Lymfocyty totiž vykazují vysokou homogenitu cytoplazmy, bez jakýchkoliv granul a s konstantním barevným podáním v celé ploše cytoplazmy. **Bazofilní granulocyty** jsou díky tmavé a masivní granulaci výrazně kontrastně odděleny od všech struktur, a to přispívá k celkovému zvýraznění bazofilů v obraze. **Monocyty** i přes jejich vysokou jasovou nehomogenitu cytoplazmy jsou zřetelně viditelné. **Eozinofilní granulocyty** jsou i ve zvýrazněném obraze zobrazeny oproti ostatním buněčným liniím nejtmavěji. Je to dáno červenou barvou cytoplazmy (barevně podobné s erytrocyty). Kvůli přítomnosti větších granul se mohou ve zvýrazněné cytoplazmě vyskytovat otvory (Obr. 25, Obr. 26). **Neutrofilní granulocyty** mají v závislosti na jemné granulaci a v některých případech prosvětleným místům nehomogenně zvýrazněnou cytoplazmu. I přes tyto nedostatky je tento připravený obraz vhodný pro následnou segmentaci. Výsledky správně zvýrazněných buněk jsou uvedeny na Obr. 26. Po podílu ve vztahu 5 vznikají hodnoty větší jak 1, proto se u nich provede přepočet na hodnotu 1.



Obr. 26 výsledky procesu zvýraznění cytoplazmy

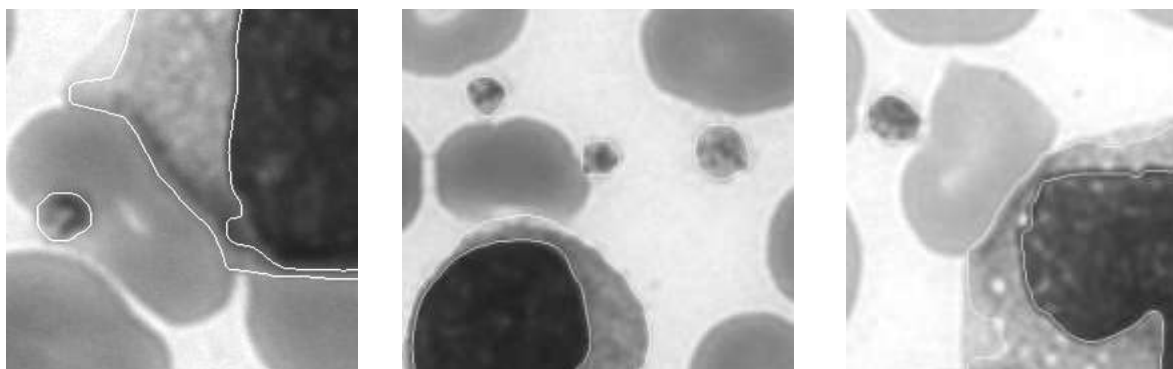
Vlastní segmentace cytoplazmy

Segmentace cytoplazmy probíhá stejným způsobem jako segmentace jádra. Před segmentováním je spočítán práh, který je důležitý pro oddělení tmavého pozadí a zbytků erytrocytů. Výpočet prahu popisuje vztah 6, ve kterém se jedná pouze o součet násobku standardní odchylky s pevně zvolenou konstantou. Tato konstanta byla stanovena na základě hodnoty zbytkových částí erytrocytů. Stanovený práh poslouží jako rozhodovací kritérium pro tvorbu binární masky metodou narůstání regionů, jak tomu bylo v případě segmentace jádra.

$$THR_{cytoplazma} = 0,1 + 1,2 \cdot \sigma(ERY) \quad (6)$$

Po dokončení segmentace cytoplazmy přichází na řadu vyčištění binární masky od nežádoucích fragmentů (krevní destičky, nespojené části cytoplazmy, zbytky erytrocytů). K tomuto účelu jsou použity morfologické operace pro *vyplnění otvorů* a odstranění drobných

regionů kruhovou maskou o průměru 9 pixelů ($0,67\ \mu\text{m}$) a odstranění velkých fragmentů nesouvisejících s jádrem funkcí, která je nastavena, aby eliminovala plochy menší než 25 000 pixelů (odvozeno z průměru kruhu 178 pixelů = $13,2\ \mu\text{m}$). Takto stanovená mez pomůže k eliminaci segmentovaných trombocytů (Obr. 27) a zároveň svojí velikostí neohrozí segmentovanou cytoplazmu leukocytu.

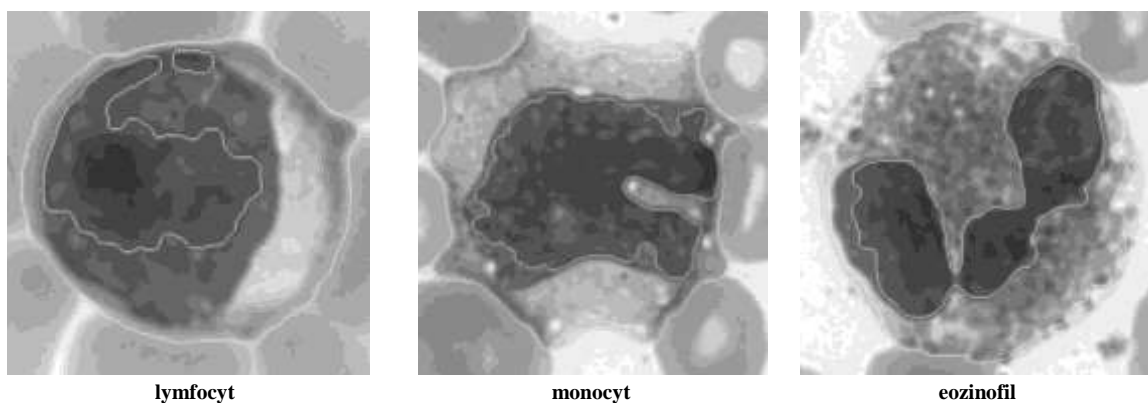


Obr. 27 případ nežádoucí segmentace trombocytů v případě nepoužití funkce pro odstranění větších regionů

5.4.6 Chybovost segmentace jádra a cytoplazmy

Nedostatečná segmentace jádra

K nedostatečné segmentaci jádra může dojít několika mechanismy. Důvodem je nejčastěji špatný odhad prahu. Nejčastěji jsou k tomuto jevu náchylné větší lymfocyty a monocyty (Obr. 28). Tyto buňky obsahují velké jádro, které má mnohdy nekompaktní chromatin plstěného charakteru. U eozinofilů dochází k těmto chybám zřídka. Jednou z možných příčin může být rozostření obrazu nebo pokles jasu.

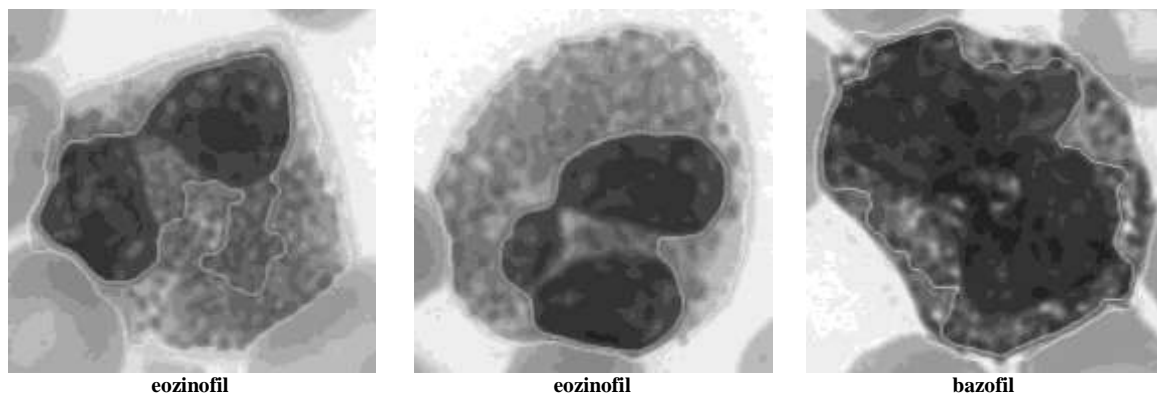


Obr. 28 příklad nedostatečně segmentovaných jaderných jader (tmavé ohraničené oblasti)

Pokud byla fotografie pořízena s barevnou odchylkou nebo nátěr není dobře probarven, tak se pravděpodobnost výskytu této chyby zvyšuje. Tato chyba pak vnáší další chyby do procesu měření dat, ve kterém falešně snižuje průměrnou barevnost cytoplazmy. Uvedený lymfocyt je příkladem extrémní chyby segmentačního algoritmu a takto nedostatečně segmentovaný se vyskytl pouze v tomto uvedeném případě.

Nadměrná segmentace jádra

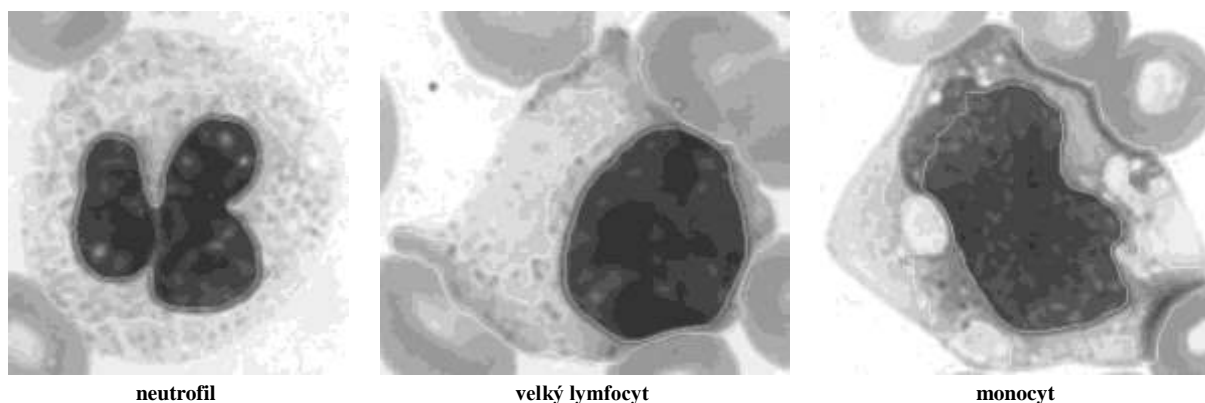
Touto chybou jsou nejčastěji postiženy obrazy buněk s tmavší granulární cytoplazmou (eozinofily a bazofily), když výpočet prahu selhává díky tmavým oblastem okolo jádra (Obr. 29). Granula jsou často nahuštěna do úzkých prostor mezi jadernými segmenty. U bazofilů je jistým jevem, že jádro celé překrývají masivní tmavá granula. Tato chyba není tak závažná, jelikož se z jádra nezískávají významné informace o barvě pro klasifikaci. Naopak je tomu u bazofilů, které tak mají falešně zvýšený obvod, plochu a poměr ploch jádra a cytoplazmy, což umožňuje jejich rozpoznání oproti neutrofilům.



Obr. 29 příklad nadměrně segmentovaných jader (vnitřní ohraničené oblasti)

Nedostatečná segmentace cytoplazmy

Tato odchylka se objevuje poměrně hojně u leukocytů se světlým typem cytoplazmy. Postihuje neutrofilů, monocytů a některé velké lymfocyty. U neutrofilů a eozinofilů bývá tato chyba způsobena nerovnoměrným rozložením granul (Obr. 30). Cytoplazma je na některých místech tmavší, ale na příklad u okrajů je příliš světlá na to, aby prahovací algoritmus zareagoval. Neutrofilů mohou být také hypogranulárního charakteru. To znamená, že takovéto neutrofilní granulocyty mají nedostatek granul a jejich cytoplazma dosahuje hodně světlých odstínů. Monocyty postihuje tato odchylka hlavně u okrajů cytoplazmy anebo okolo vakuol. Vzhledem k velkému objemu buňky a její ploše se u mnoha buněk objevují v cytoplazmě nerovnoměrně prosvětlená místa. Eozinofily zde nejsou uvedeny.



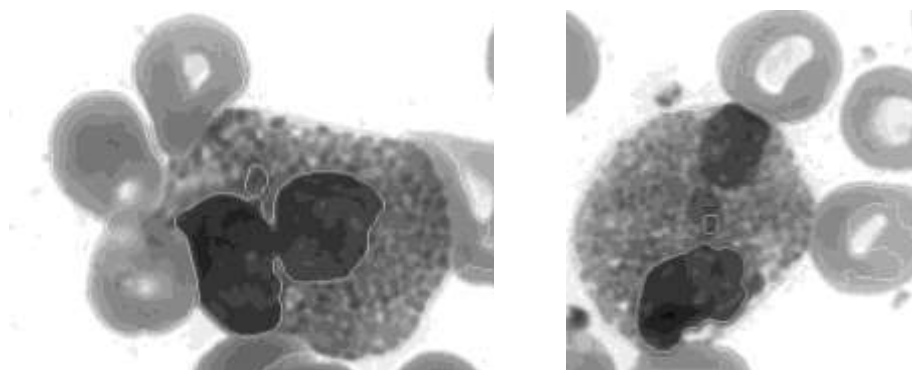
Obr. 30 příklad nadměrně segmentované cytoplazmy (světlé oblasti)

Pro měření geometrických parametrů se cytoplazma jeví jako málo použitelná. Avšak u malých lymfocytů, kterých je nadbytek oproti velkým, může být dobrým rozpoznávacím znakem od monocytů. Cytoplazma je důležitá pro její barevnost a její výhodou je, že když je nedostatečně segmentována, tak přesto i malý počet pixelů podá v průměru dostatečně přesnou informaci pro další klasifikaci. Na základě vysoké barevné variability je velmi obtížné navrhnout univerzální segmentační funkci pro všechny buněčné typy. Proto je segmentace cytoplazmy brána spíše jako odhad[9].

Nadměrná segmentace cytoplazmy

Nadměrná segmentace cytoplazmy se projevuje přechodem masky převážně do naléhajících erytrocytů. Výskyt této chyby není tak častý, byl zaznamenán jen u buněk s výraznou barevnou odchylkou (Obr. 31). Erytrocyty jsou díky jinému barevnému podání snímku tmavší a negativně tak ovlivňují výpočet prahu. V některých případech jsou leukocyty zaznamenány v místech nátěru, kde nejsou erytrocyty rovnoměrně rozprostřeny a mohou být přes sebe přeskládány. Pokud jsou erytrocyty takto rozloženy, dochází k jejich ztmavení.

Těmto chybám lze předejít nasnímáním fotografií za konstantních podmínek. Všeobecným pravidlem pro zaznamenávání fotografií z krevních nátěrů je výběr vrstvy s rovnoměrným rozložením buněk. Tato oblast se často označuje jako „monolayer“.



Obr. 31 příklad nadměrně segmentované cytoplazmy eozinofilů (oblasti mimo buňku)

5.5 Příznaky oblasti jádra a cytoplazmy

Po segmentaci jádra a cytoplazmy je nutné se zaměřit na sběr možných dat, která budou sloužit pro následné určení typu buňky. Jak již bylo zmíněno ve výše uvedeném literárním průzkumu [10; 14; 15], lze získat několik druhů parametrů z obou struktur (jádro a cytoplazma). Na základě literárních poznatků byly pro tuto práci zvoleny **geometrické a intenzitní příznaky**.

Mezi geometrické příznaky byly pro účely analýzy vybrány:

- **obvod** jádra a cytoplazmy,
- **plocha** jádra a cytoplazmy,

- **poměr obvodů a poměr ploch** jádra a cytoplazmy.

Intenzitní příznaky:

- **průměrná intenzita** oblasti jádra v R, G a B kanálu,
- **průměrná intenzita** oblasti cytoplazmy v R, G a B kanálu,
- **šíkmost histogramu** průměrných intenzit oblasti jádra v R, G a B kanálu,
- **špičatost histogramu** průměrných intenzit oblasti jádra v R, G a B kanálu.

Obvod jádra i cytoplazmy byly spočítány pomocí součtu pixelů v binárním obraze vytvořeném hranovým operátorem. Plocha cytoplazmy a jádra vychází z binárních obrazů metody narůstání regionů, kde je proveden součet pixelů binární masky zastupující daný objekt. Tyto dvě masky slouží také k výpočtu průměrné intenzity z R, G a B složek obrazu (jádro a cytoplazma). Výpočet je proveden s pixely v originálním obraze odpovídajícími pixelům hodnoty 1 v příslušné binární masce. Do souboru parametrů byly vybrány i parametry histogramu související s intenzitními složkami jádra a cytoplazmy. Jedná se o šíkmost a špičatost histogramu a byly zařazeny k intenzitním příznakům.

5.5.1 Analýza dat jádra a cytoplazmy

Naměřená data byla zpracována ve statistickém softwaru STATISCICA 10. Celkem bylo pro účely analýzy dat použito 330 buněk nasnímaných ze 17 různých nátěrů (viz Tab. 7). U všech příznaků byl proveden základní statistický průzkum a testy normality. Počty jednotlivých populací jsou poměrně rozmanité, což vyplývá z jejich nerovnoměrného zastoupení v periferní krvi.

Hlavní zájem byl směřován především na příznaky barevnosti jádra a cytoplazmy, kde již od pohledu jsou znatelné výrazné odstínové difference. Získané geometrické příznaky jsou pouze druhořadými faktory. Na základě níže uvedených zjištění byly některé buňky z analýzy vyřazeny. Poslední sloupec Tab. 7 obsahuje počty buněk získaných z vyřazených fotografií a nově nasnímaných buněk. Tato poslední skupina figuruje až v níže uvedené části o klasifikaci neuronových sítí a zde je uvedena pouze informativně.

Buňky	získané	stav po vyřazení	nové pro testování
Neutrofily	92	71	33
Lymfocyty	96	88	31
Monocyty	56	45	26
Eozinofily	68	55	15
Bazofily	18	15	6
Celkem	330	276	111

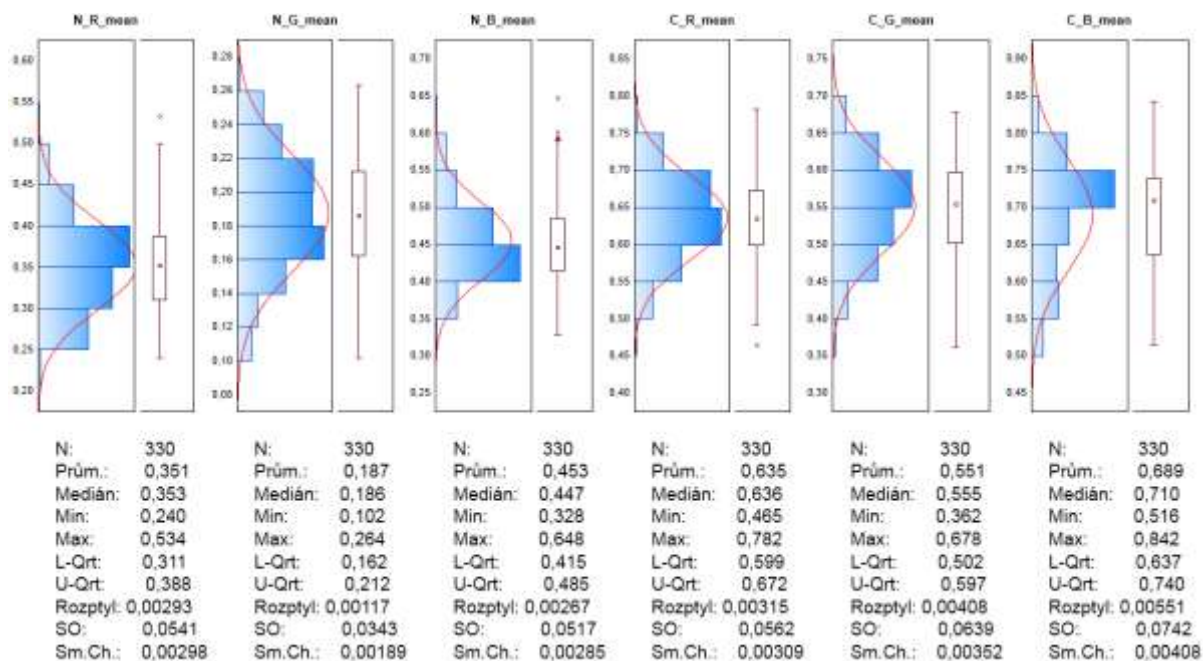
Tab. 7 počty získaných buněk

5.5.2 Statistický souhrn příznaků

Barevné vlastnosti jádra a cytoplazmy

Segmentační algoritmus provedl na základě vyčleněných oblastí výpočty průměrů intenzit v jednotlivých barevných kanálech RGB obrazu. Intenzity jader jsou vesměs tmavšího charakteru. Graf 2 (N_R_mean, N_G_mean, N_B_mean) ukazuje, že zelená složka jádra je nejtmaší a modrá složka nejsvětější. Odstín jádra se odvíjí od jeho zralosti a stádia chromatinu, který navazuje barvivo. Jádro může do jisté míry přejímat odstín cytoplazmy, která pod ním leží, a tím posunout intenzitu v barevných složkách obrazu do vyšších hodnot (eozinofily, monocyty).

Grafy průměrů intenzit barev cytoplazmy (Graf 2, C_R_mean, C_G_mean, C_B_mean) mají oproti jádru širší rozsah a u modré složky nerovnoměrné rozložení. To je důsledek barevné rozmanitosti cytoplazmy jednotlivých buněčných typů. Z grafu je také patrné, že rozsahy intenzit jsou podobné, ale s mírně odlišným umístěním ukazatele středu.



Graf 2 statistický přehled průměrů intenzit barevných složek (R, G, B) jádra (N) a cytoplazmy (C)

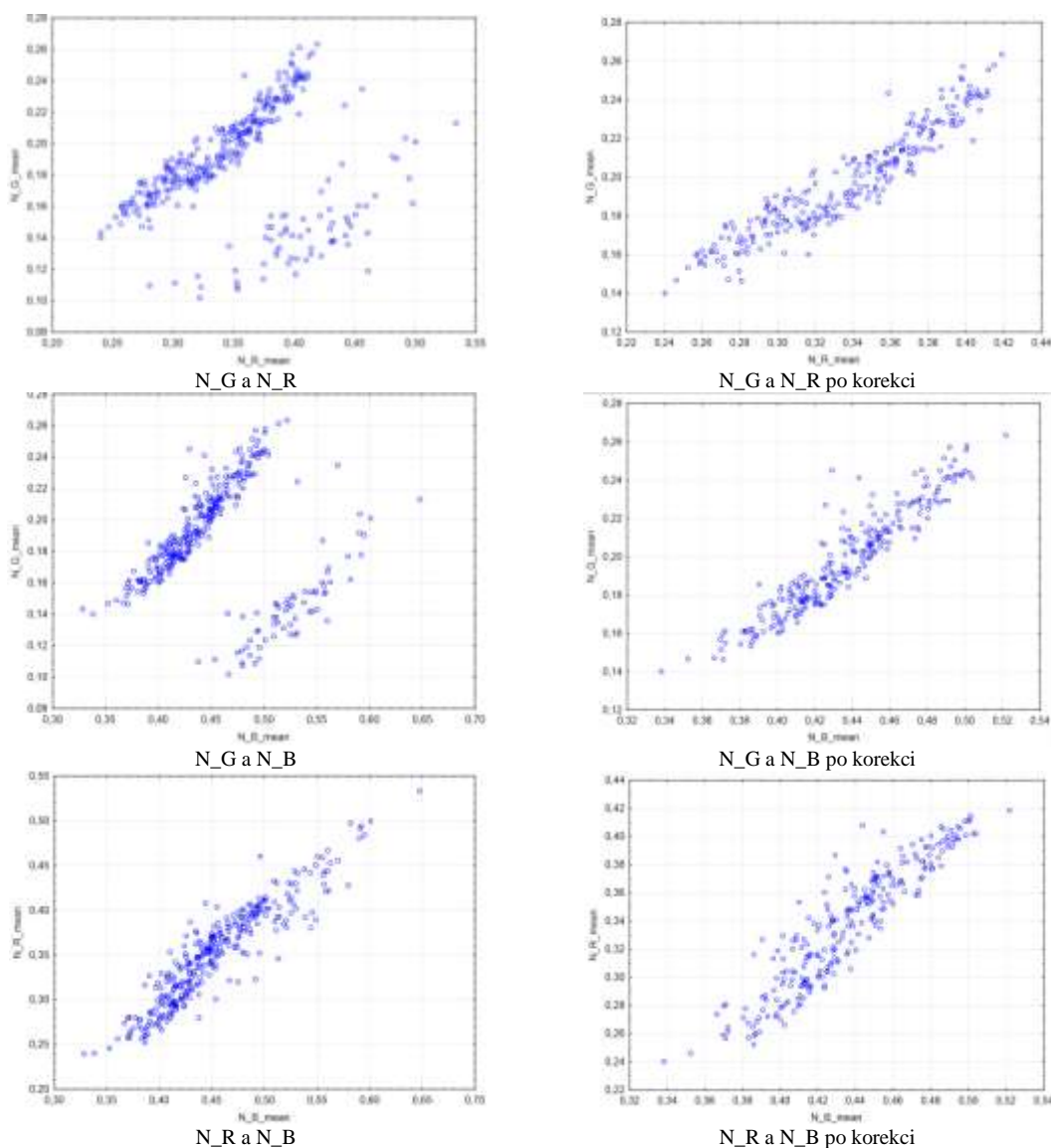
Čištění souboru dat

Saphiro-Wilksovy testy spolu s normálními p-grafy (Příloha 2) potvrdily shodu intenzitních příznaků s normálním rozložením. Tyto příznaky jsou vhodné pro další zpracování a nevyžadují další výpočetní korekce. Odlehle hodnoty je lepší v tomto případě vyloučit pro zachování jednotnosti souboru.

Další krok v analýze příznaků zahrnuje zjištění jejich vztahů mezi populacemi leukocytů. Náležitý přehled podávají korelační bodové grafy, ve kterých jsou proti sobě vytyčeny potřebné příznaky. Bodové grafy (Graf 3) popisují intenzitní příznaky jádra. Tato

skupina grafů barevné intenzity jádra poukazuje na odlehlou množinu buněk od více početného shluku. V odlehlé množině jsou zastoupeny všechny druhy buněk. Dle lineárního posunu (nejvíce patrné u zelené složky), lze usuzovat, že se jedná o odchylku vzniklou při pořizování fotografií. Tato chyba vznikla pravděpodobně při nastavování clony kondenzoru mikroskopu, jelikož nebyla přítomna stupnice a aretační pozice. Další možností je nedostatečná soustřednost osvětlení mikroskopu s objektivem.

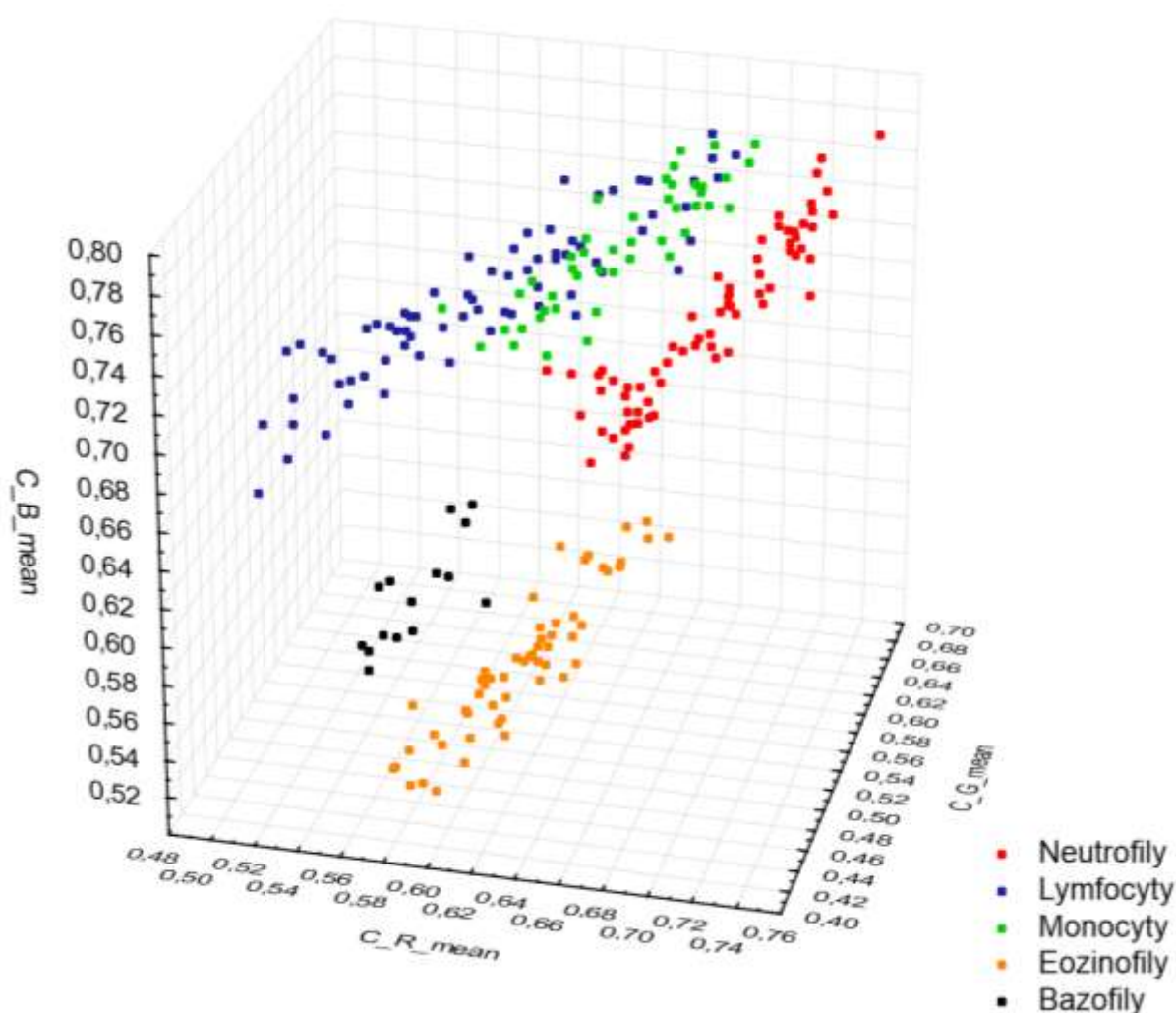
Buňky s touto chybou byly vyřazeny z dalších analýz v rámci přesnějšího vymezení populací v prostoru. Zpětným dohledáním původu buněk bylo zjištěno, že odchýlené buňky patří do několika (3–4) náterrů snímáných v totožném období. Graf 3 také popisuje výraznou podobnost zastoupených příznaků (vysoký korelační koeficient), proto tyto faktory nenesou žádný význam pro klasifikaci. Zmíněné příznaky barevnosti jádra významně přispěly pouze ke zjištění a eliminaci chybových dat.



Graf 3 intenzity barevných kanálů (R, G, B) oblasti jádra (N), vlevo: zdvojený shluk (barevná odch.), vpravo: po korekci

Popsaná barevná odchylka zaznamenaná u oblasti jádra se také projevila v grafech barevné intenzity cytoplazmy (0). V tomto případě se jednalo pouze o mírný rozptyl původního shluku. Po odstranění buněk v případě barevnosti jádra se shluky lépe oddělily.

Graf 4 znázorňuje vztah mezi třemi příznaky (intenzita cytoplazmy R, G a B složky obrazu) po vyřazení nevhodných buněk (původní graf: Příloha 7). Celkem jsou v grafu znatelné 4 významné shluky. *Tento graf potvrzuje původní domněnku* o výrazné barevné odlišnosti buněčných populací na základě barvy cytoplazmy (viz kapitola 5.3.2). Pro lepší představu byly vytvořeny také 2D grafy popisující Graf 4 (Příloha 8).



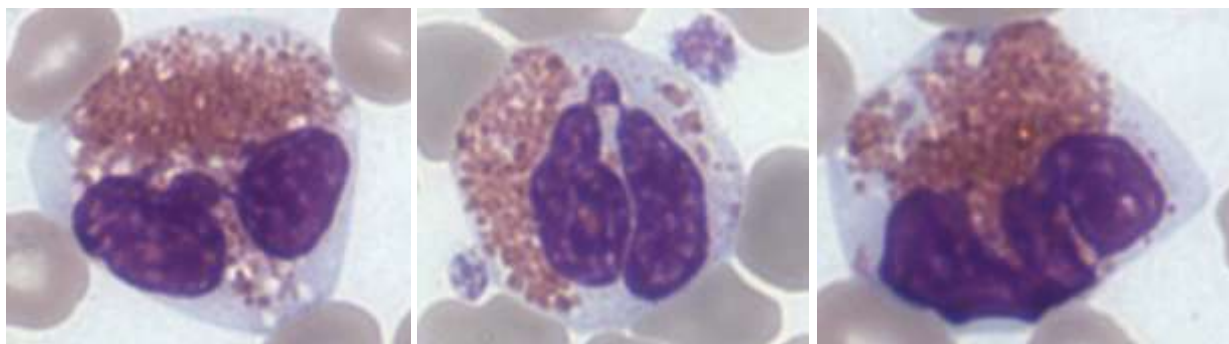
Graf 4 intenzita barevných kanálů (R, G, B) cytoplazmy (C)

Shluky jsou zřetelně lineárně oddělitelné, zejména **eozinofily**, které jsou díky cihlově červené cytoplazmě nejvíce posunuté do červeného pásma. V původním grafu (Příloha 7) jsou eozinofily více rozprostřeny (barevná odchylka), ale také některé z nich zasahují blíže k populaci neutrofilů. Důsledkem této odchylky je vzhled fotografovaného eozinofilu, který má nehomogenní rozložení cytoplazmy (Obr. 32). Mezi červenými granulami jsou značně velké modré prostory posouvající barevnost cytoplazmy.

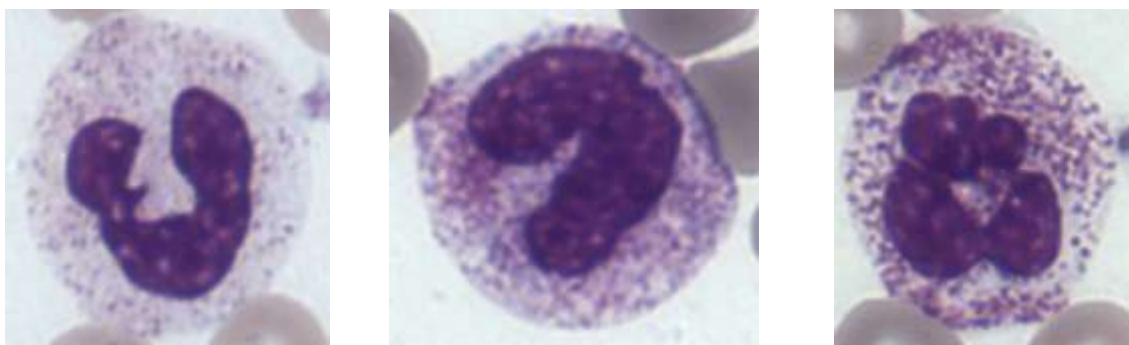
Bazofily jsou rovněž jasně odděleny od ostatních buněčných populací, avšak jejich nízký počet vnáší jistou chybu pro popis shluku.

Neutrofily vytvářejí rovnoměrně rozložený oddělený shluk, ale v grafu bez korekce (0) tento shluk zasahuje do shluku bazofilů. Tento fakt nemusí být zapříčiněn pouze zmiňovanou barevnou odchylkou, ale i přítomností toxické granulace neutrofilů (Obr. 33), která má sytější tóny granul podobných bazofilním granulocytům.

Lymfocyty a monocyty jsou od ostatních populací umístěny nejvýše, avšak mezi sebou tvoří společný shluk a nelze je v tomto případě dobře oddělit.



Obr. 32 eozinofily s nerovnoměrným rozložením cytoplazmy



Obr. 33 neutrofily s toxickou granulací (vlevo fyziologický neutrofil)

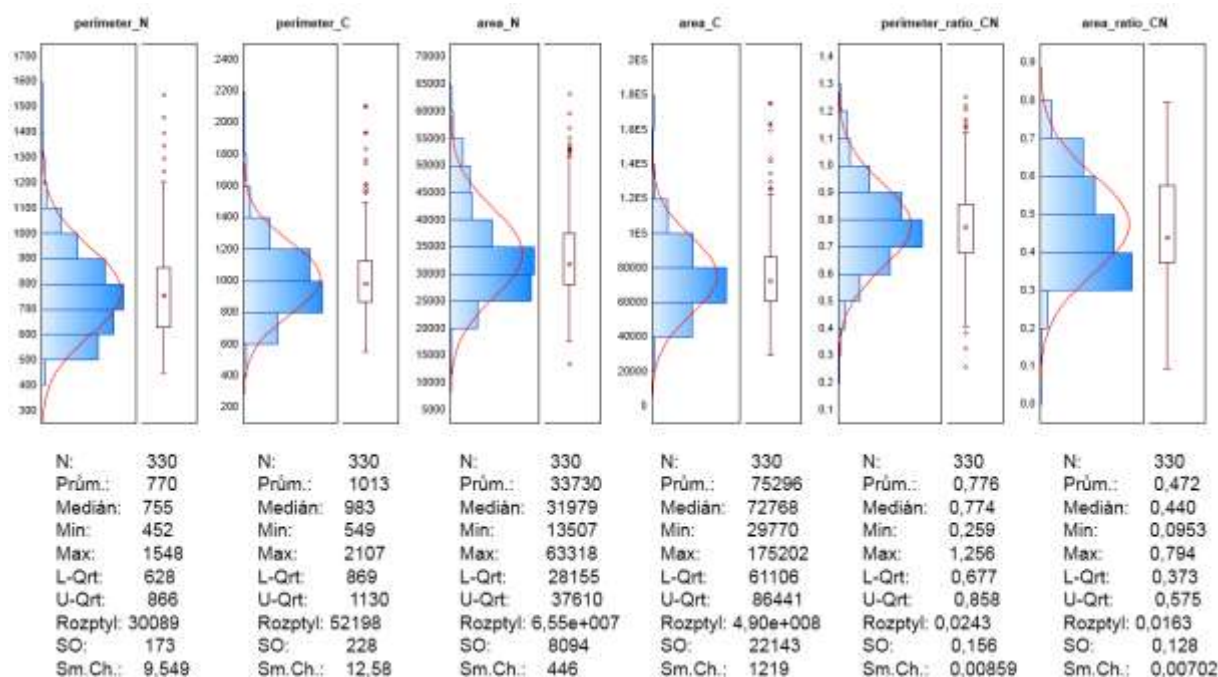
Geometrické příznaky jádra a cytoplazmy

Pro lepší přehled byly nejprve zpracovány všechna nasbíraná data a následně provedeno jejich pročištění. Jak ukazuje Graf 5, téměř všechny geometrické příznaky vykazují normální rozložení s několika odlehlými a extrémními hodnotami. V některých případech se může stát, že buňka má abnormální rozměry či bizarní tvar. V této skupině jsou výrazně zastoupeny zejména monocyty. Další chybovost může být způsobena segmentačním algoritmem, který ne vždy dokáže zcela správně odseparovat buněčné části.

U segmentace jader jsou problematické hlavně linie granulocytů, nejvíce eozinofily a bazofily, jelikož přítomná masivní granulace překrývá jádro. Přítomnost granul nahuštěných na povrch jádra může způsobovat také lokální ztmavení obrazu ve všech barevných složkách. Rozložení jsou kompaktní, bez výrazného členění do více skupin. Jelikož je množství

použitých buněk nesouměrné, je jejich vliv na histogram bezpředmětný. I přes tyto nedostatky mohou geometrické proporce posloužit pro dodatečné rozeznání dvou populací mezi sebou.

Histogram plochy jádra a cytoplazmy (Graf 5) vypovídá o tom, že nejfrekventovanější jsou buňky s plochou přibližně 30 000 pixelů (plocha jádra) a 60 000 (plocha cytoplazmy, odpovídající poloměr 20,5 μm). Uvedená čísla nejčastěji přísluší granulocytům a lymfocytům, které mají v součtu největší zástup a nízký rozptyl.

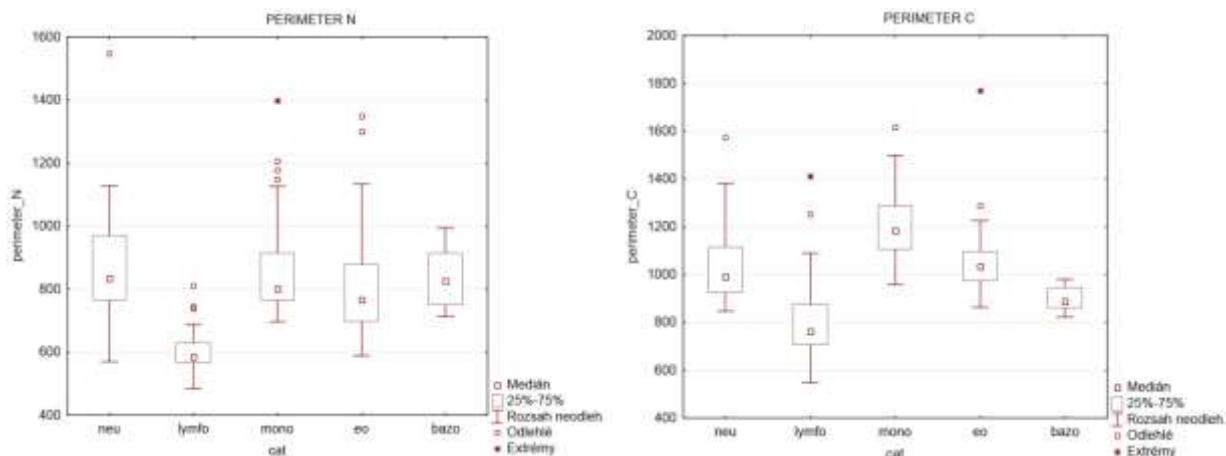


Graf 5 statistický přehled geometrických příznaků jádra (N) a cytoplazmy (C) v pixelech a jejich poměrů

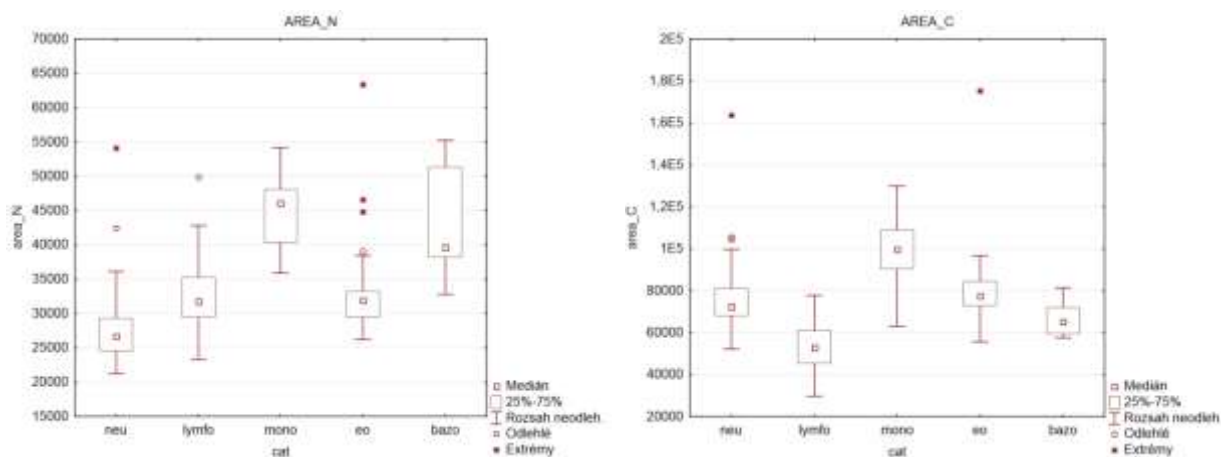
Níže uvedené 3 grafy znázorňují rozložení populací v každém geometrickém příznaku. Pro rozlišení dvou populací buněk je vhodným kritériem obvod jádra a cytoplazmy (Graf 6). Graf obvodu jádra zobrazuje, že kromě lymfocytů mají všechny populace podobný rozsah a jejich rozdělení leží na stejné úrovni. Populace leukocytů v grafu oblasti cytoplazmy jsou lépe odděleny. Granulocyty (neutrofily, eozinofily a bazofily) všeobecně nejsou vhodné pro geometrické rozdělení, protože jejich velikosti jsou téměř stejné a jejich rozptyly se překrývají, což znemožňuje jejich rozpoznání. Granulocyty mají také variabilní jádro rozdělené do segmentů a tím pádem mohou nepřesně vykazovat větší obvod jádra. Zvláštní význam mají tyto obvody pro monocyty a lymfocyty, protože jejich velikosti jsou odlišné a oddělitelné.

Graf 7 popisuje statistické rozložení plochy jádra a cytoplazmy. Porovnáním jednotlivých populací je vidět že neutrofily, lymfocyty a eozinofily se překrývají. Výjimkou u plochy jádra jsou bazofily, které jsou výše postavené. Je to dáno tím, že plochu jádra negativně ovlivňuje masivní granulace. Do tohoto rozmezí zasahují i některé velké lymfocyty. Monocyty jsou umístěny na úrovni bazofilů a nepodílejí se tolik na výše uvedené hodnotě ploch jádra a cytoplazmy.

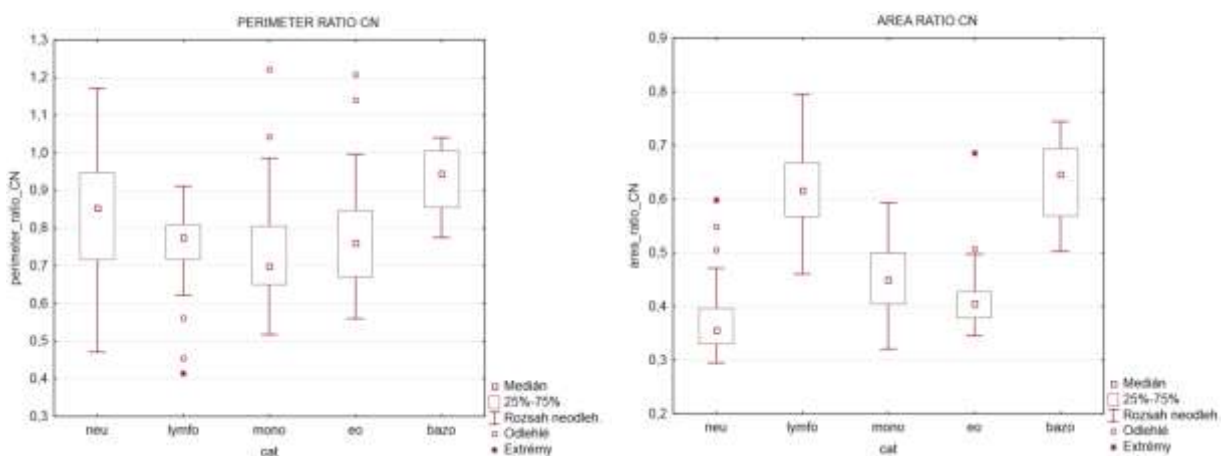
U ploch cytoplazmy (Graf 7 vpravo) je znatelné překrytí v rozložení jednotlivých populací. Všechny granulocyty mají přibližně stejnou velikost cytoplazmy. Graf také dokazuje, že nejmenšími buňkami jsou lymfocyty. Na základě tohoto parametru mohou být dobře rozlišitelné monocyty a lymfocyty.



Graf 6 zleva: obvod jádra (N) a cytoplazmy (C) pro každou buněčnou linii zvlášť



Graf 7 zleva: plocha jádra (N) a cytoplazmy (C) pro každou buněčnou linii zvlášť



Graf 8 vlevo: poměr obvodů cytoplazmy a jádra, vpravo: poměr ploch cytoplazmy a jádra (všechny populace)

Poměr obvodů a poměr ploch cytoplazmy a jádra poskytuje dobrý přehled o celkovém vzhledu buňky (Graf 8). Poměr vyjadřuje jak je jádro velké vzhledem k cytoplazmě. Pokud je daný poměr nízký, znamená to, že buňka má co do plochy více cytoplazmy a menší jádro (monocyty, některé neutrofilní granulocyty, větší lymfocyty). V opačném případě vysoký poměr poukazuje na malý obsah cytoplazmy, často jen úzký lem okolo jádra, jaký je typický pro malé lymfocyty nebo bazofily.

V ojedinělých případech mohou být extrémní hodnoty známkou brzké nebo prodloužené činnosti segmentačního algoritmu cytoplazmy. Výhodou geometrických proporcí buněk je nezávislost na případném barevném posunu v intenzitě barev jakékoliv barevné složky jádra i cytoplazmy.

Testy normality (Příloha 3) prokázaly významnou shodu naměřených dat s normálním rozdělením. Data jsou vhodná pro další zpracování, avšak je potřeba vzít v úvahu odlehle hodnoty, které mohou v konečném řešení zhoršit možnou klasifikaci buněk.

Šikmost a špičatost histogramu cytoplazmy

Pro rozšíření souboru příznaků bylo uvažováno i využití prvků charakterizujících tvar histogramu, korespondující s jasovým zastoupením pixelů v obraze. Základní statistiky ukázaly, že šikmost (kurtosis) a špičatost (skewness) histogramu nejsou vhodně interpretovatelné. Saphiro-Wilksův test potvrdil, že šikmost není shodná s normálním rozdělením (Příloha 4). Špičatost histogramu se ve dvou případech ze tří shodovala s normálním rozdělením (S_R – červená složka, S_B – modrá složka), v jednom případě byla méně podobná (S_G – modrá složka). Korelační graf těchto parametrů (0) zobrazuje převahu korelovaných vazeb mezi jednotlivými příznaky. Šikmost histogramu vykazovala zřetelnější rozložení shluků, ale po bližším prozkoumání nebyl mezi rozprostřenými skupinami leukocytů shledán větší význam. Pro další manipulaci by hodnoty šikmosti vyžadovaly normalizaci či přepočet. Po dalším pohledu na data se mezi jednotlivými skupinami buněk neprojevily žádné významné odlišnosti. Vzhledem k množství významnějších faktorů již v jiných postupech není s těmito příznaky počítáno.

5.5.3 Výběr příznaků

Statistickým průzkumem byly vybrány pro klasifikaci příznaky, které nejlépe charakterizují a odlišují jednotlivé buněčné frakce. Jako významné příznaky byly potvrzeny proměnné nesoucí informaci o barevnosti cytoplazmy (C_R_mean, C_G_mean, C_B_mean), které jsou schopny odlišit až 4 třídy leukocytů. Doplňkovými příznaky byly určeny geometrické informace (Area_N, Area_C, Area_ratio_CN, Perimeter_C), které mohou umožnit přesnější rozpoznání některých tříd (neutrofily/bazofily, lymfocyty/monocyty).

5.6 Programová klasifikace příznaků

Poslední fáze algoritmu pro rozpoznání struktur v obraze a získání jejich příznaků je jejich klasifikace. Vzhledem k výpočetní nenáročnosti, snadnější adaptivitě a většímu počtu příznaků byla zvolena klasifikace pomocí neuronových sítí. Tato metoda byla realizována v programovém prostředí MATLAB, resp. pomocí několika funkcí Neural Network Toolboxu. Architekturou sítě byla zvolena dopředná síť s algoritmem učení založeným na zpětném šíření chyby (*Feed-forward backpropagation*)[16].

5.6.1 Topologie sítě

Prvotní návrh

Původním záměrem bylo vytvořit jednotnou neuronovou síť, která by dokázala hodnotit všechny zastoupené třídy leukocytů na základě dvou nekorelovaných příznaků. Takto nastavená síť měla 2 vstupy (C_R_mean a C_B_mean – intenzitní příznaky cytoplazmy) a pět výstupů. Po naučení se však ukázalo, že síť má nestabilní chybovost a je tudíž nevhodná. Problematický byl rovněž velmi nízký počet bazofilů oproti ostatním populacím, což snižovalo citlivost sítě na tuto buněčnou linii. Některé další populace, které často mívají nízkou homogenitu cytoplazmy (monocyty, velké lymfocyty), byly nesprávně řazeny k jiným cílovým skupinám. Při zvyšování počtu skrytých neuronů se sice stabilita chybovosti mírně zvýšila, ale hrozilo přílišné specializování sítě na daný soubor. Tímto krokem by byla znemožněna generalizace sítě pro další použití na jiných fotografiích. Z tohoto důvodu byl tento postup vyloučen a již nebude zmiňován.

Sloučení tříd leukocytů

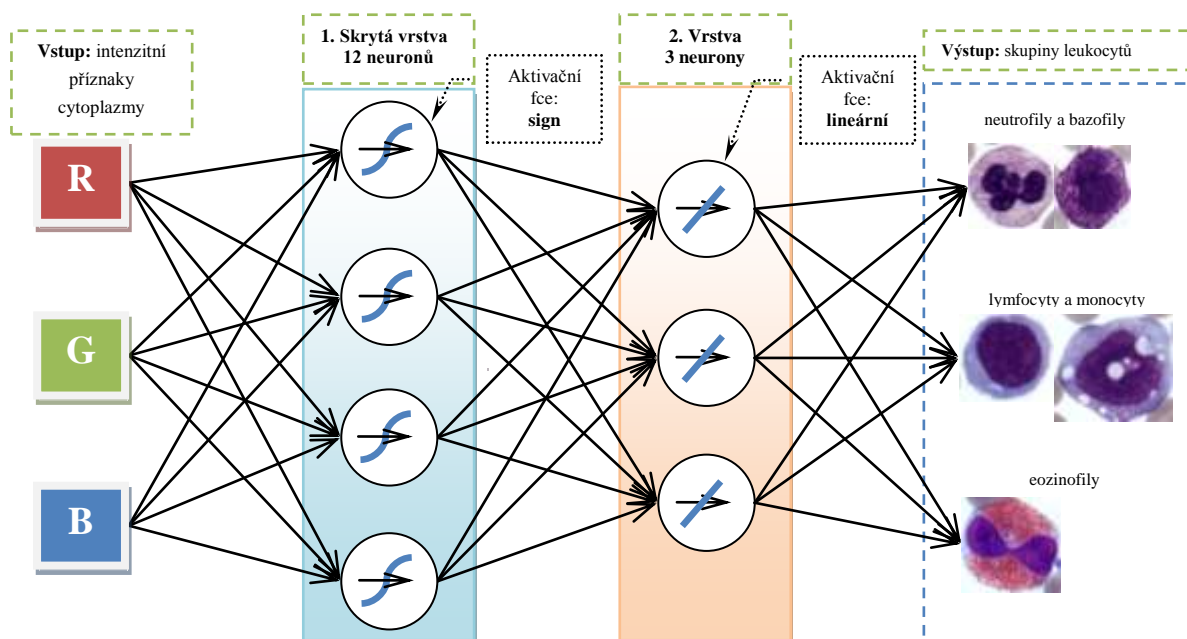
Vhodnějším prostředkem jak od sebe cílové skupiny lépe odlišit je využít neuronové sítě v několika etapách. Graf 4 poukazuje na významnou barevnou podobnost mezi cytoplazmou lymfocytů a monocytů. Tento fakt umožňuje sloučení těchto dvou buněčných linií do jedné a tuto společnou skupinu pak hodnotit separátním klasifikátorem.

Druhým řešeným problémem se stal nízký počet bazofilních granulocytů. Nasnímané bazofily se svým počtem 15 buněk nebyly dostatečně reprezentativním vzorkem pro efektivní rozlišení od více početných skupin leukocytů. Jak již bylo zjištěno, jsou bazofily na stejné barevné linii spolu s neutrofily (Graf 4), pouze s posunem k tmavším odstínům. Z tohoto důvodu byly tyto dvě populace sloučeny pro další určovací proces.

Tímto snížením počtu výstupních skupin bylo docíleno přesnějšího určení skupin. Pro zajištění stability sítě a snížení chybovosti byl k původním dvěma příznakům (R a B složka cytoplazmy) dodán třetí příznak (G složka cytoplazmy). Geometrické příznaky v tomto případě použity nebyly, jelikož nebyly dostatečně reprezentativní pro rozlišení leukocytů.

Hlavní neuronová síť

V závislosti na sloučení skupin byla navržena dvouvrstvá síť se vstupem pro tři příznaky a výstupem do tří skupin (viz Obr. 34). Jako vstup slouží naměřené příznaky barevných kanálů cytoplazmy, které jsou propojeny se všemi neurony první vrstvy. První (skrytá) vrstva obsahuje 12 neuronů. Názory na počet neuronů ve skryté vrstvě se často rozcházejí a doporučují se dle různých výpočtů. Není však jasně stanovené pravidlo, které by mělo jednoznačně určit počet skrytých neuronů[16]. Počet neuronů je závislý na využití sítě a mnohdy je nutné jej stanovit experimentálně na základě řešeného problému a oscilace chyb vniklých při opakovaném učení sítě.



Obr. 34 struktura navržené hlavní neuronové sítě (schématicky)

Jednou z nejčastěji používaných aktivačních (transformačních) funkcí je funkce signum [16], v tomto případě byla použita funkce `tansig` (tangenciální sigmoida), (Obr. 35 vlevo). Druhá (výstupní) vrstva obsahuje stejný počet neuronů jako výstupních skupin. Neurony druhé vrstvy jsou vybaveny lineární transformační funkcí bez dalších úprav (omezení rozsahu) a slouží k vlastní klasifikaci předložených příznaků. V případě lineární transformační funkce byla použita lineární funkce `purelin` (Obr. 35 vpravo).

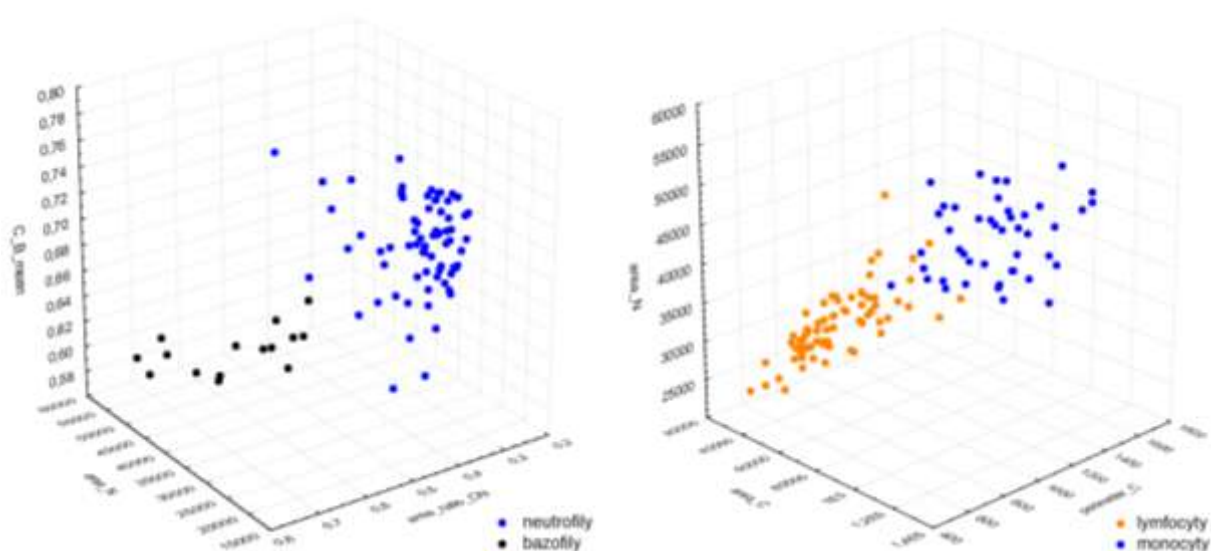


Obr. 35 transformační funkce, vlevo `signum` (skryté neurony), vpravo lineární transf. funkce (neurony 2. vrstvy)

Na výstupu této sítě se generuje vektor o délce shodné s počtem skupin a obsahuje hodnoty v rozsahu, který má kladný i záporný charakter. Tyto hodnoty vyjadřují *stupeň příslušnosti* klasifikované buňky k výstupní skupině. Stupněm příslušnosti je myšlena vyjádřená hodnota, která charakterizuje podobnost klasifikované buňky s cílovou skupinou. Čím blíže k 1 je hodnota ve vektoru, tím buňka nabývá vyššího stupně příslušnosti k dané skupině.

Vedlejší neuronové sítě

Stejným uvedeným způsobem dle Obr. 34 jsou navrženy další dvě neuronové sítě pro následnou klasifikaci **neutrofilů/bazofilů** a **lymfocytů/monocytů**. Tyto dvě sítě mají stejný počet vstupů, ale s jinými příznaky. Pro diagnostiku neutrofilů/bazofilů byly zvoleny geometrické příznaky: plocha jádra (Area_N), poměr ploch cytoplazmy a jádra (Area_ratioCN) a intenzitní příznak: modrá složka cytoplazmy (C_B_mean). K rozlišení lymfocytů/monocytů jsou použity pouze geometrické příznaky: plocha jádra (Area_N), plocha cytoplazmy (Area_C) a obvod cytoplazmy (Perimeter_C). Skrytá vrstva obsahuje jen 10 neuronů oproti původním 12. Druhá vrstva má pouze 2 výstupní neurony. Klasifikace pro skupinu neutrofilů a bazofilů je spuštěna tehdy, pokud je daná buňka hlavní sítí označena jako první skupina. Stejný postup s jinými vstupy je použit u skupiny lymfocytů a monocytů. Graf 9 zobrazuje rozložení sloučených populací buněk v prostoru v závislosti na zvolených příznacích (soubor naměřených dat). Tento graf poskytuje představu o prostoru, podle kterého budou klasifikovat neuronové sítě.



Graf 9 rozdělení buněčných řad v prostoru na základě zvolených příznaků

Použité funkce

Ve fázích zkoušení a testování byly sítě nejdříve tvořeny pomocí GUI manažeru z Matlab Neural Network toolboxu. Tímto byla odzkoušena obecná schopnost vytvořit fungující síť. Pro použití klasifikace buněk a dosažení lepší kontroly nad průběhem manipulace se sítěmi bylo potřeba některé parametry sítě měnit a z tohoto důvodu bylo využito návrhu sítě ve skriptu pomocí předpřipravených funkcí zmíněného toolboxu. Především se jedná o funkce: `newff`, `net`, `trainlm` a `sim`.

Funkce `newff` slouží pro vytvoření sítě se zvoleným počtem neuronů. Funkce `net` vybavuje již vytvořenou síť transformačními funkcemi dle výběru. Funkce `trainlm` slouží k samotnému trénování vytvořené sítě na předložených vstupech Levenberg-Marquardtovým algoritmem (viz níže). Funkce `sim` slouží pro simulaci naučené sítě na předloženém libovolném souboru s příznaky. Tato funkce je zakomponována do klasifikačního algoritmu a její výstup v podobě vektoru příslušnosti (viz výše) je konečným výsledkem k určení druhu leukocytu.

5.6.2 Trénování sítě

Aby bylo docíleno lepších výsledků, snížení chybovosti a zjištění chování sítě pro jednotlivé skupiny, je nutné síť uvést do trénovacího (učebního) procesu. Pro tyto účely slouží tzv. trénovací množina, která se skládá ze dvou stejně dlouhých vektorů. První vektor obsahuje vzorové příznaky (řádky) pro potřebný počet cílových skupin (sloupce). Vzorové příznaky do sítě vstupují v takovém tvaru, v jakém byly předešlým algoritmem získány. V druhém vektoru se nachází informace o skupině, ke které každý soubor příznaků (sloupec) patří. Na základě znalosti vztahu cílových skupin a souboru příznaků lze síť opakovaně učit, aby bylo docíleno generalizace.

Zpětné šíření chyby

Zpětné šíření chyby (*backpropagation*) je využíváno jako iterační proces, během kterého se síť opakovaně učí na stejném souboru příznaků a upravuje své váhy dle výsledků. Výběr vzorů lze realizovat sekvenčně, nebo náhodně. V případě této práce jsou příznaky voleny náhodným výběrem. Doba, kdy dojde k vystřídání všech prvků v souboru minimálně jednou, se nazývá *epocha*[16]. K učení je zapotřebí několikanásobně proces zopakovat do momentu, kdy chyba učení (globální – střední kvadratická odchylka) klesne na minimum nebo pod zvolenou hranici. K zastavení procesu učení může sloužit i maximální stanovený počet epoch.

Na výstupu každé epochy se počítá energie (pomocí energetické funkce), která označuje odchylku mezi reálnými a požadovanými hodnotami výstupů neuronové sítě. V každém cyklu dochází k optimalizaci energetické funkce tak, aby měla klesající tendenci.

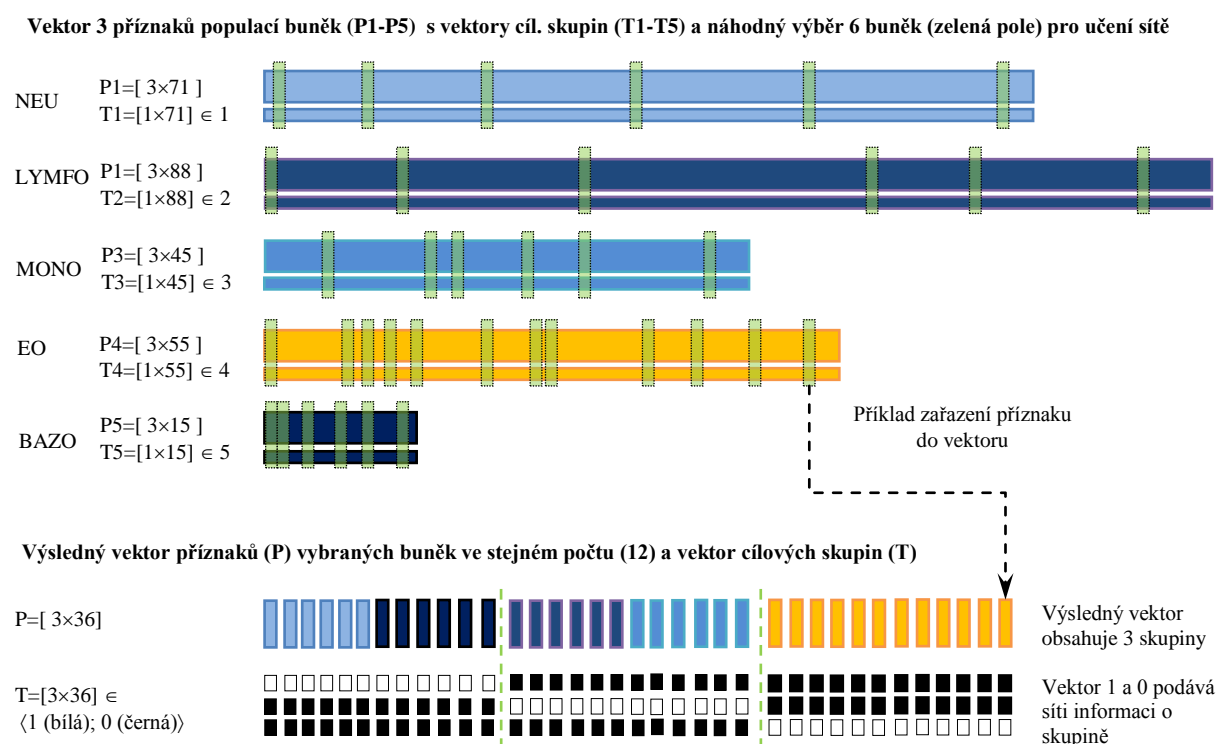
Optimalizace energetické funkce je využívána pro celkové uzpůsobení sítě pro trénovací množinu s důrazem na nalezení nejvhodnější aproximace mezi jednotlivými vzory.[16]

Levenberg-Marquardtova optimalizace

Jako optimalizační proces je v případě této práce využívána Levenberg-Marquardtova optimalizace[16]. Tato metoda hledá učící pravidla při zachování krátkých kroků během strmého poklesu energetické funkce, nebo dlouhých kroků při mírném poklesu energetické funkce (časová úspora, zrychlení algoritmu). Jedná se o metodu gradientního postupu vycházející z Gauss-Newtonovy metody[16].

Generování testovací množiny

V rané fázi testování byl postup uzpůsoben tak, aby byl z každé skupiny vybrán stejný počet reprezentativních vzorků a přitom zůstaly i vzorky na zpětné otestování. Losování bylo navrženo náhodně. Pomocí funkce `randi`, doplněné o sadu podmínek aby se již vylosovaná čísla neopakovaly, byl vygenerován vektor čísel. Každé číslo vektoru symbolizuje vylosovanou pozici v množině příznaků jedné skupiny buněk. Tyto vektory pro jednotlivé buněčné řady jsou pak spojeny do jednoho a tvoří tak zkušební vektor příznaků (Obr. 36).



Obr. 36 příklad tvoření učebního vektoru 12 buněk z každé populace leukocytů pro hlavní neuronovou síť

Zbýlý vektor příznaků po vylosování (vyjmutí) učební množiny pak slouží jako zkušební množina, která nezasahuje do procesu učení. U této zkušební množiny se pak vypočítá chybovost klasifikace (počet chybných/celkový počet). Trénovací algoritmus předloženou skupinu rozdělí v poměru 80 : 20. Pro vlastní trénování slouží 80 % prvků

trénovací množiny a zbylých 20 % pro validaci. Validační proces nezasahuje do učebního procesu, ale slouží jako odezva trénování na nastavení vah.

5.6.3 Chybovost sítě

Výše uvedený postup se osvědčil na výběru přibližně 30 buněk z každé množiny, kdy chybovost oscilovala přibližně okolo 18 %. Jako ideální počet neuronů ve skryté vrstvě se osvědčilo 12 neuronů u hlavní sítě a 10 neuronů pro obě dvě vedlejší sítě. Při snížení počtu neuronů se zvýšila oscilace chyby klasifikace.

Další možnou variantou pro snížení chybovosti klasifikace bylo zvětšení trénovací množiny. Oproti předchozímu postupu se změnil počet vybraných buněk na 80 % ze všech buněčných populací. Tímto krokem se snížila chybovost klasifikace přibližně na 10 %. V rámci zpřesnění generalizace sítě byla nakonec pro učení sítě využita celá množina dat. Toto rozhodnutí bylo učiněno z důvodu dalšího využití sítě. Výsledky chybovosti ukazuje Tab. 8. Chybovost se v tomto případě snížila a oscilovala mezi 5–10 %.

Trénovací poměr	Hlavní síť (3 třídy)		Vedlejší síť - lymfocyty a monocyty		Vedlejší síť - neutrofily a bazofily	
	80 % první skupiny	100 % první skupiny	80 % první skupiny	100 % první skupiny	80 % první skupiny	100 % první skupiny
Chyba	5,5 %	3,81 %	3,7 %	9 %	6 %	5,1 %
Standardní och.	0,4	1,1	2,8	0,75	1,5	4

Tab. 8 tabulka procentuálních chyb pro dva učební poměry (testovací množiny, 111 buněk)

Rozšíření učební množiny na celkový počet buněk jistým způsobem zohledňuje nerovnoměrný počet populací leukocytů v periferní krvi. Jelikož je neutrofilních granulocytů v krvi nejvíce (až 70 %), je třeba, aby neuronová síť spolehlivě konvergovala k tomuto druhu. To samé platí pro lymfocyty a monocyty, které jsou přítomny v periferní krvi v množství okolo 10 %. Eozinofily a bazofily mají procentuální zastoupení nejnižší (jednotky procent). Jelikož byla celá populace buněk použita pro učební proces, byla pro klasifikaci sestavena nová populace buněk (také uvedena jako „testovací skupina“). Celkový počet buněk je 111 a obsahuje různorodou směs všech zastoupených linií.

Zajímavý je fakt, že testovací skupina buněk téměř z poloviny obsahuje buňky vyřazené kvůli barevným odchylkám (viz Graf 3). Tyto buňky byly pro učební proces nevhodné, protože by znemožňovaly přesnou aproximaci sítě prostorem příznaků. Nic ovšem nebrání tomu síť použít pro otestování jejího chodu i na horších, neideálních (vyřazených) datech. Druhou polovinu testovací skupiny tvoří nové, dodatečně nasnímané fotografie, které by měly sloužit k ověření reprodukovatelnosti. To znamená, zda je síť schopna pracovat na nově dodaných, korektních datech.

Výše zmíněné odstavce popisují chybovost jednotlivých sítí bez jakéhokoliv vzájemného vlivu. Lze předpokládat zvýšení chybovosti, jelikož chyby ze dvou sériově zapojených sítí se mohou násobit. Proto byla skupina testovacích buněk prověřena na plné konstrukci klasifikačního systému. Jak je vidět v Tab. 9, senzitivita a specificita jsou poměrně vysoké.

Buňky	neutrofily	lymfocyty	monocyty	eozinofily	bazofily
Celkový počet	31	29	26	15	6
Správně určené	29	27	23	15	4
Falešně negativní	2	2	3	0	2
Falešně pozitivní	1	3	1	1	3
senzitivita	93,9 %	93,5 %	89,7 %	100 %	75,0 %
specificita	96,9 %	90,6 %	96,3 %	93,8 %	66,7 %
Úspěšnost klasifikace	102 ze 111 buněk = 91,9 % (8,1% chyba)				

Tab. 9 výsledky kompletní klasifikace třemi neuronovými sítěmi

Neutrofily se vyznačují vyváženou senzitivitou a specificitou. **Lymfocyty a monocyty** mají také oba parametry vyvážené. Výsledky ukázaly, že monocyty v několika případech mohou být mylně klasifikovány jako lymfocyty, což dokazuje nižší senzitivita sítě. Naopak vyšší specificita pro monocyty značí, že prakticky žádný lymfocyt nebyl zařazen k monocytům. **Eozinofily** mají nejlepší výsledky ze všech skupin, lze to připisovat specifické barvě cytoplazmy (odlehlý shluk) a také tomu, že eozinofily jako jediné buňky není potřeba klasifikovat druhou neuronovou sítí.

Nejhorší výsledky vzhledem k malému souboru a větším odchylkám při separaci jádra a cytoplazmy podaly **bazofily**. Do jisté míry se na této chybě projevila dysbalance v počtu neutrofilů a bazofilů při učení sítě, protože síť lépe konverguje k více zastoupené množině.

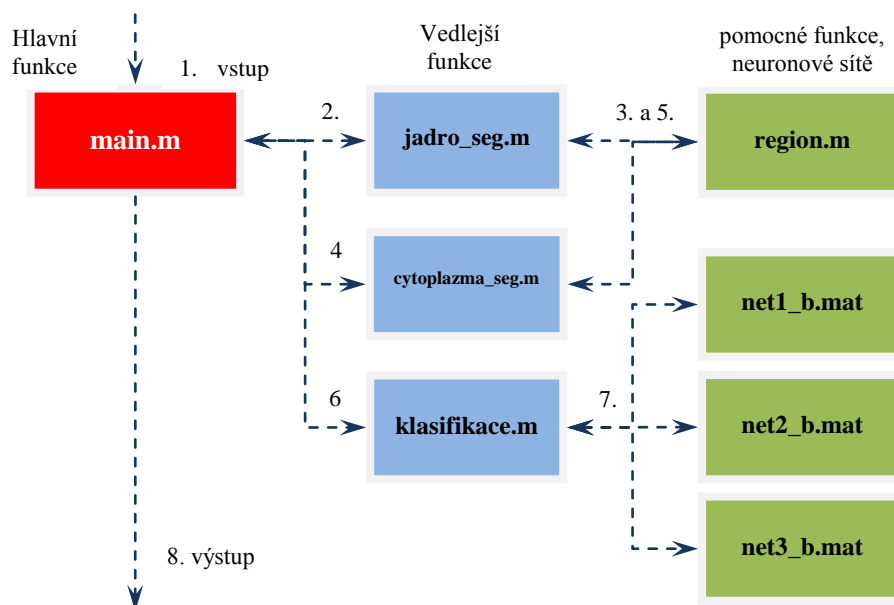
počet	buňky	zařazené do	populace
2 a 1	neu	↔	bazo
1	lymfo	→	bazo
3 a 1	lymfo	↔	mono
1	bazo	→	eo

Tab. 10 počet buněk chybně zařazených do jiných populací, šipky značí kam byly buňky nesprávně zařazeny

Tab. 10 vyjadřuje vztah migrace buněk mezi cílovými skupinami. Podle počtu je vidět, že hlavní síť dokáže leukocyty rozřadit správně (7 z 9) a větší chybovost vzniká až na úrovni vedlejších sítí vždy mezi dvěma skupinami.

6 Rozvržení algoritmu v programovém prostředí MATLAB

Navržený proces byl celkově rozložen do 5 funkcí navržených v programovém prostředí MATLAB (Obr. 37). Jedná se o hlavní funkci `main.m` a ostatní, vedlejší či pomocné funkce. Na Obr. 37 zobrazuje sled volání jednotlivých funkcí.



Obr. 37 struktura celého algoritmu navrženého v MATLABu

6.1 Funkce `main.m`

Tato funkce je hlavní funkcí. Slouží pro zavedení vstupního obrazu, který předkládá ostatním funkcím pro zpracování. Do této funkce se vrací binární obrazy segmentovaného jádra a cytoplazmy vypočtené vedlejšími funkcemi (`jadro_seg.m` a `cytoplazma_seg.m`). Ze získaných proměnných ze segmentačního procesu jsou zpracovány příznaky a poté jsou předloženy funkci `klasifikace.m`. Výstupem funkce `main.m` je informace o typu buňky a soubor vykreslených obrazů (vstupy, výstupy).

6.2 Funkce `jadro_seg.m` a `cytoplazma_seg.m`

Funkce `jadro_seg.m` a `cytoplazma_seg.m` jsou vedeny jako funkce vedlejší. Tyto segmentační funkce pracují dle navrženého schématu (Obr. 19). Nejdříve je obraz předzpracován a pak následuje samotná segmentace metodou *narůstání regionů*. Pro tento účel je volána pomocná funkce `region.m`, která provede metodu *narůstání regionů*. Funkce

`jadro_seg.m` i `cytoplazma.m` vrací do hlavní funkce `main.m` binární masku segmentovaného jádra nebo cytoplazmy a binární masku s obvodem jádra nebo cytoplazmy.

6.2.1 Funkce `region.m`

Funkce `region.m` slouží pro realizaci metody narůstání regionů na předložených binárních obrazech ze segmentace cytoplazmy nebo jádra. Dalším vstupem je originální obraz, který slouží pro porovnání intenzitních hodnot pixelů s prahem. Funkce obsahuje rozhodovací podmínku, která na základě vstupující funkce realizuje k ní určenou metodu narůstání regionů. Technicky se nepodařilo realizovat všeobecnou funkci pro narůstání regionů z důvodu požadavku na práh. Pro segmentaci jádra je určena funkce, která zapíše pixel hodnoty 1 do binární masky, pokud je pixel originálního obrazu menší než hodnota předloženého prahu. U segmentační funkce cytoplazmy rozhoduje o zapsání hodnota větší než hodnota prahu.

6.3 Funkce `klasifikace.m` a neuronové sítě

Klasifikační funkci `klasifikace.m` je předložen soubor příznaků. Nejprve je volána první/hlavní neuronová síť (`net1_b.mat`), která vyhodnotí 3 skupiny leukocytů (neutrofily/bazofily, monocyty/lymfocyty a eozinofily). Pokud je buňka určena jako eozinofilní funkce `klasifikace.m` končí a předloží informaci funkci `main.m`. Pokud je však vyhodnocena jedna ze skupin leukocytů, je volána jedna z funkcí pro dodatečné rozpoznání na základě jiných příznaků. Skupinu monocyty/lymfocyty zpracovává funkce `net2_b.mat` a skupinu neutrofily/bazofily zpracovává funkce `net3_b.mat`. Vyhodnocené třídy jsou vráceny pomocí klasifikační funkce do hlavní funkce.

6.3.1 Neuronové sítě `net1_b.mat`, `net2_b.mat` a `net3_b.mat`

Každá z těchto neuronových sítí je realizována jako soubor funkcí, vytvořených pomocí Matlab Neural Network toolboxu. Obsahují informaci o počtu vrstev, neuronů a typu aktivačních funkcí. Tyto neuronové sítě byly naučeny na souboru příznaků (viz kap. 5.6.2).

7 Diskuze

7.1 Barevné odchylky

Stěžejním a nejčastěji zmiňovaným aspektem této práce je kvalita obrazu a zajištění konstantních snímacích podmínek. I přes snahu jednotného nastavení mikroskopu a snímací aparatury nebyly tyto podmínky vždy ideální. Snímání leukocytů probíhalo v několika etapách v závislosti na čase a vytíženosti laboratoře. Jelikož na používaném mikroskopu probíhaly praktická cvičení, bylo nutné před každým snímáním provést kalibraci osvětlení a celkové očištění přístroje.

Jedním z faktorů přispívajícím na odchylku intenzity a kontrastu je uvažováno nastavení clony kondenzoru. Tato clona nemá fixovatelný ovládací prvek a během manipulace s mikroskopem se může uvolnit a pozměnit tak snímací podmínky. Na ovládacím prvku kondenzorové clony není přítomna stupnice, podle které by šlo přesné nastavení clony monitorovat.

7.2 Různé počty leukocytů

Samotnou odchylku ve snímání může způsobovat krevní nátěr, který může mít odlišný barevný nádech pozadí kvůli přítomným bílkovinám. Pokud jsou leukocyty snímány v části, kde jsou rovnoměrně rozmístěny (monolayer), jsou fotografie a buňky vhodně zobrazeny. V opačném případě může vzniknout obraz s barevným posunem nebo jasovou odchylkou v důsledku nakupení erytrocytů v několika vrstvách, nebo utlačením leukocytu a zvýšením sytosti jejich cytoplazmy.

Dalším důležitým bodem diskuze je počet buněk jednotlivých buněčných populací pro účely trénování neuronové sítě. Každá z buněčných populací má odlišné procentuální zastoupení v periferní krvi. Proto bylo některých buněk (bazofily, eozinofily) získáno mnohem méně než ostatních. Pro dosažení lepších výsledků klasifikace by bylo ideální získat od každé populace stejný počet buněk ve větším počtu. Tím by bylo docíleno vyšší senzitivity a specifity klasifikace.

7.3 Důraz na výběr vzorových buněk

Po zhodnocení chybovosti jak segmentačních algoritmů, tak klasifikačního procesu byly chybné buňky zpětně zrevidovány a bylo zjištěno, že v některých případech mají drobné fyziologické odchylky. Mezi tyto odchylky patří: toxická granulace neutrofilů, nehomogenní rozmístění granul eozinofilů, odlišná barevnost monocytů (Obr. 28 a Obr. 29). V těchto několika případech může segmentační a klasifikační algoritmus chybovat. Jako prevence by pouze pro účely trénování bylo vhodné zaměřit se na ryze fyziologické buňky anebo provést důkladnější rozdělení buněk do podskupin. Jelikož jsou buňky jedinečnými elementy

s vysokou variabilitou vzhledu, je ovšem každé další škálování obtížné, často bez náležitého efektu.

7.4 Možné rozšíření práce

Tento odstavec je uveden jako návrh rozšíření této práce pro případnou návaznost na toto téma. Možným rozšířením by mohlo být nalezení optimálních podmínek pro snímání buněk z mikroskopu a kompenzace barevných odchylek. Na základě takto zajištěného procesu záznamu dat se pak zaměřit na zdokonalení segmentačních procesů a rozšíření souboru příznaků. Dalším možným směrem by mohlo být eventuální rozšíření souboru klasifikovaných krvinek, a to na cílenější specifikaci nebo na raná stádia (extrémně nízký výskyt v periferní krvi). Například lymfocyty lze dále dělit na velké a malé. Granulocytárním segmentům předchází vývojové stádium tyče. Některé patologické stavy by bylo možné také klasifikovat na základě vzhledu buněk. Například u granulocytů lze mikroskopicky rozlišit toxickou granulaci, abnormální velikost buněk, hyposegmentaci, hypersegmentaci, hypogranulaci nebo hypergranulaci.

Jistým řešením pro aplikaci této práce do praxe by mohlo být vytvoření automatizovaného procesu segmentace. Víze spočívá v napojení snímací aparatury přímo na mikroskop, který je manuálně ovládán. Obsluha mikroskopu vloží krevní nátěr a zajistí jeho promítnutí na kameru. Dále obsluha manuálně vyhledá vhodnou oblast pro fotografování (monolayer), dle hustoty a rozložení erytrocytů by bylo teoreticky možné získat na základě navrženého algoritmu. Po nalezení libovolné bílé krvinky by obsluha provedla zaostření, které by bylo monitorováno na základě změny kontrastu v obraze. Konečným procesem by bylo nasnímání obrazu a jeho klasifikace. Takto klasifikované buňky by se zaznamenávaly do databáze pro daný krevní nátěr a bylo by možné provést korekci a výpočet zastoupení jednotlivých leukocytů v periferní krvi.

V této práci je uvažován pouze případ jednoho leukocytu v obraze. v některých případech může nastat situace, kdy leukocyty na sebe naléhají a tento problém by měl být řešen tak, aby algoritmus rozpoznal jednotlivé buňky a klasifikoval každou zvlášť. Fotografie, které byly pořízeny z mikroskopu, jsou rozsáhlé a na jejich ploše se běžně vyskytuje několik leukocytů (Příloha 9). Automatickým nalezením a oříznutím buněk v zaznamenaném obraze by odpadla fáze ručního předzpracování. Poté by algoritmus byl schopen hodnotit více buněk v jednom obraze.

8 Závěr

Tato práce popisuje hematologickou problematiku spojenou se zpracováním obrazů krevních nátěrů. První kapitola práce popisuje laboratorní metodu zvanou „Diferenciální počet leukocytů“. Dále jsou popsány jednotlivé formy bílých krvinek, které jsou touto metodou hodnoceny. Další část je tvořena popisem významných prací, zabývajících se analýzou a klasifikací pomocí digitálního zpracování obrazu. Kvůli časté podobnosti metod byly uvedeny jen některé důležité aplikace zabývající se segmentací obrazu, nebo jeho následnou klasifikací.

Pro budoucí softwarové zpracování bylo nasnímáno (mikroskop Nikon Eclipse 80i, kamera Nikon DS-Fi1) přibližně 450 fotografií bílých krvinek ve světlém poli. Zároveň bylo experimentálně pořízeno 350 duplicitních buněk fázovým kontrastem. Ke snímání bylo použito 20 krevních nátěrů. Do souboru bylo zařazeno 5 typů leukocytů, fyziologicky se vyskytujících v periferní krvi. I když byl kladen důraz na zajištění standardních snímacích podmínek, tak v několika případech došlo k mírným barevným odchylkám. Získané fotografie jsou přiloženy na digitálním médiu a obsahují jak originální fotografie seřazené dle nátěrů, tak fotografie předpřipravené ořezem a roztříděné dle buněčného typu. Všechny fotografie jsou uloženy bez jakýchkoliv úprav.

V praktické části byla věnována pozornost základním vlastnostem pořízených snímků. Součástí této kapitoly je rozbor struktur v obraze barveného krevního nátěru. Hlavní část zde tvoří statistické zpracování buněčných součástí na bázi geometrických a barevných vlastností buněk. Na základě statistické analýzy jsou extrahované hodnoty použity pro segmentaci jádra a cytoplazmy. Segmentace jádra je principiálně postavena na prahování obrazu a metodě narůstání regionů. Segmentace cytoplazmy využívá nově navrženého předzpracování obrazu, postaveného na sérii matematických úprav mezi jednotlivými barevnými kanály. Jedna z kapitol je věnována výsledkům těchto dvou segmentačních postupů. Segmentované jádro a cytoplazma jsou tímto způsobem připraveny pro měření příznaků a klasifikační proces. Algoritmus byl realizován v programovém prostředí MATLAB.

Po naměření potřebných dat ze segmentačního algoritmu byl proveden statistický průzkum za účelem nalezení reprezentativních příznaků pro klasifikační proces. Nejlépe popisujícími příznaky byly stanoveny barevné složky cytoplazmy doplněné geometrickými příznaky.

Jako klasifikační funkce byla zvolena metoda umělých neuronových sítí. Celkový návrh byl rozdělen do tří neuronových sítí s účelem rozložit potenciál pro hodnocení jednotlivých buněčných populací. Síť byly naučeny na předložených parametrech leukocytů získaných z výstupu segmentačního algoritmu. První síť klasifikuje leukocyty do tří tříd, z nichž dvě jsou sloučené (neutrofily/bazofily a lymfocyty/monocyty). Sloučené populace

leukocytů jsou klasifikovány podle specifických příznaků jinou neuronovou sítí. Rozdělení klasifikačního procesu do tří sítí pomohlo zvýšit jejich stabilitu a zvýšit přesnost klasifikace.

Klasifikační algoritmus navržený v programovém prostředí MATLAB byl otestován na 111 snímcích, které byly z poloviny tvořeny daty nevhodnými pro učení sítě a druhou polovinou regulérními fotografiemi. I přes tyto okolnosti bylo dosaženo 91,9% úspěšnosti klasifikace. Senzitivita a specificita se u čtyř populací buněk (neutrofily, lymfocyty, monocyty, eozinofily) pohybuje okolo hranice 90 % a u jedné populace (bazofily) okolo 75 % (senzitivita) a 67 % (specificita).

9 Slovník pojmů

- **Morfologie** - vědní obor biologie, který se zabývá vnější stavbou organismů. Od fyziologie se tedy liší tím, že zkoumá tvar a nikoliv funkci.
- **Segmentace jádra** – Jádro postupem vyztváření mění svůj tvar, až je tvořeno menšími částmi vzájemně propojenými tenkými jadernými můstky.
- **Bazofilie cytoplazmy** – cytoplazma se barví bazofilními barvivy (dofialova).
- **Eozinofilie cytoplazmy** – cytoplazma se barví oxyfilními barvivy (dočervena).
- **Neutrofilní cytoplazma** – cytoplazma zbarvená domodra.
- **Efektorové buňky** – krátce žijící buňky účastníci se imunitní reakce, bojující s daným antigenem.
- **Barvení May Grünwald – Giemsa Romanowski** – barvicí technika:
fixujeme roztokem May Grünwald a barvíme roztokem Giemsa-Romanowski.
Základem tohoto barvení jsou:
 - a) Kationtové (zásadité) barvivo azur B: váže se na aniontové části molekul a barví modrošedě nukleové kyseliny (DNA nebo RNA), nukleoproteiny, granule bazofilů a sekundární granula neutrofilů.
 - b) Aniontové (kyselé) barvivo eozin Y: váže se na kationtové části molekul proteinů a barví oranžovočerveně hemoglobin a eozinofilní granula.
- **Lýza buňky** – eliminace buňky za pomoci degradace její buněčné membrány a rozpuštění jejího obsahu.

10 Citovaná literatura

- [1]. **MATÝŠKOVÁ, M.** *Hodnocení nátěru periferní krve – Pokyny pro vypracování průvodního listu.* [PDF dokument] místo neznámé: SEKK, 2009. staženo ze: www.sekk.cz.
- [2]. *Automated Digital Morphology – a big step forward in standardising the differential blood count.* [PDF document] Norderstedt, Germany: SYSMEX EUROPE GMBH, 2008. Automated Digital Morphology Sysmex Xtra Online 08/2008, firemní materiál.
- [3]. **CORNET, E.; PEROL, J.-P.; THOUSSARD, X.** Performance evaluation and relevance of the CellaVision DM96 system in routine analysis and in patients with malignant hematological diseases, *International journal of laboratory hematology*. 20. January 2007.
- [4]. **PAVLÍK, J.** *Analýza diferenciálního rozpočtu v periferní krvi různými diagnostickými metodami.* Brno: Vysoké učení technické v Brně, fakulta elektrotechniky a komunikačních technologií, 2010.
- [5]. **PECKA, M.** *Laboratorní hematologie v přehledu – Fyziologie a patofyziologie buňky.* Český Těšín: FINIDR s.r.o., 2006. ISBN 80-8682-02-1.
- [6]. **RAD, A.** wikimedia commons. <http://commons.wikimedia.org>. [Online] [Citace: 26. 12. 2009.]
<[http://upload.wikimedia.org/wikipedia/commons/6/69/Hematopoiesis_\(human\)_diagram.png](http://upload.wikimedia.org/wikipedia/commons/6/69/Hematopoiesis_(human)_diagram.png)> zdroj obrázku.
- [7]. **PAVLÍK, J.; FÁTOROVÁ, I.** *Interaktivní atlas krvinek periferní krve.* [CD-ROM] 2008.
- [8]. **SADEGHIAN, F., a další.** *A Framework for White Blood Cell Segmentation in Microscopic Blood Images Using Digital Image Processing.* [Biological Procedures Online; Biomedical and Life Sciences]. místo neznámé: BioMed Central, 1. 12. 2009. Sv. 11, 1, stránky 196-206. Dostupné z URL: <<http://dx.doi.org/10.1007/s12575-009-9011-2>>. ISSN: 1480-9222; DOI: 10.1007/s12575-009-9011-2.
- [9]. **SANPANICH, A.; IAMPA, W.; PINTAVIROOJ, C.; TOSRANON, P.** White Blood Cell Segmentation by Distance Mapping Active Contour. *IEEE Xplore*. [Online] 2008. [Citace: 25. 10. 2011.] <<http://ieeexplore.ieee.org>>. ISBN 978-1-4244-2335-4; INSPEC 10412883.
- [10]. **THEERA-UMPON, N.** White Blood Cell Segmentation and Classification in Microscopic Bone Marrow Images. [editor] L. WANG a Y. JIN. *Fuzzy Systems and*

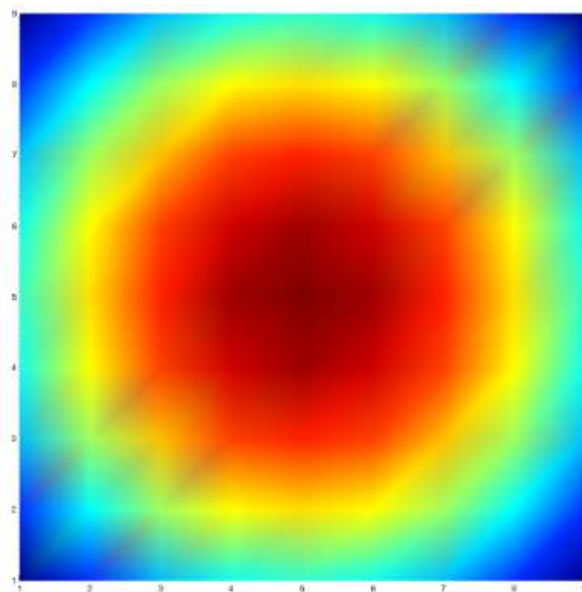
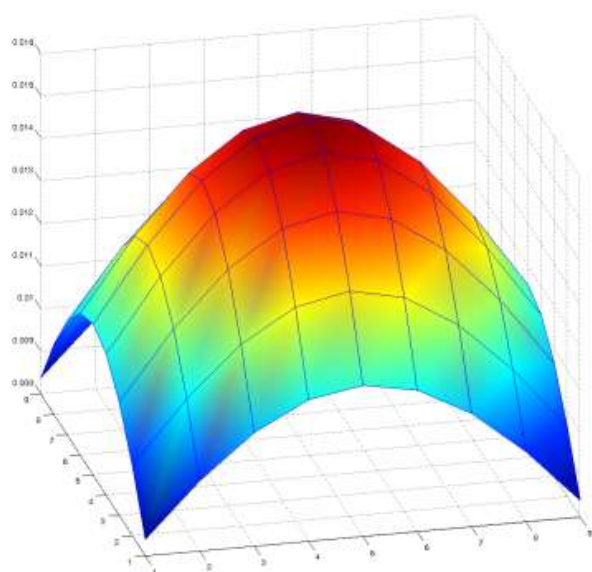
- Knowledge Discovery*. Berlin/Heidelberg: Springer Berlin/Heidelberg, 2005, Sv. 3614, stránky 787-796. Dostupné z URL: <http://dx.doi.org/10.1007/11540007_98>.
- [11]. **THEERA-UMPON, N. a DHOMPONGSA, S.** *"Morphological Granulometric Features of Nucleus in Automatic Bone Marrow White Blood Cell Classification*. místo neznámé: IEEE Xplore, 2007. Sv. 11, 3, stránky 353-359. Dostupné z URL: <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4167903&isnumber=4167888>>. DOI: 10.1109/TITB.2007.892694.
- [12]. **DORINI, L.B., MINETTO, R. a LEITE, N.J.** White blood cell segmentation using morphological operators and scale-space analysis. *SIBGRAPI 2007. Computer Graphics and Image Processing*. místo neznámé: IEEE Xplore, 2007. stránky 294-304. Dostupné z: <<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4368197&isnumber=4368154>>. DOI: 10.1109/SIBGRAPI.2007.33.
- [13]. **MARTÍNEZ-CARBALLIDO, J., ROSAS-HUERTA, C. a RAMÍREZ-CORTÉS, J.** *Metamyelocyte nucleus classification using a set of morphologic templates*. [2010 Electronics, Robotics and Automotive Mechanics Conference]. místo neznámé: IEEE Xplore, 2010. DOI: 10.1109/CERMA.2010.143.
- [14]. **PLASTININ, A., KUPRIYANOV, A. a ILYASOVA, N.** Color textural analysis of the blood preparation images. *Optical Memory & Neural Networks*. 01. 09 2008, Sv. 3., 17, stránky 201-207. ISSN: 1060-992X; DOI: 10.3103/S1060992X0803003X; Dostupné z: <<http://dx.doi.org/10.3103/S1060992X0803003X>>.
- [15]. **MARKIEWICZ, T., OSOWSKI, S. a MARIANSKA, B.** [editor] B. Beliczynski et al. *White Blood Cell Automatic Counting System Based on Support Vector Machine*. Berlin/Heidelberg: Springer-Verlag Berlin Heidelberg, 2007. 4432, stránky 318-326.
- [16]. **Pavlík, D.** Využití neuronové sítě při identifikaci znaku v obraze. [Diplomová práce]. Brno: autor neznámý, 2008. str. 62. dostupné online (8. 5. 2012) z: <http://www.vutbr.cz/www_base/zav_prace_soubor_verejne.php?file_id=7365>; vedoucí práce: Ing. Michal Kohout.

11 Doplnková literatura

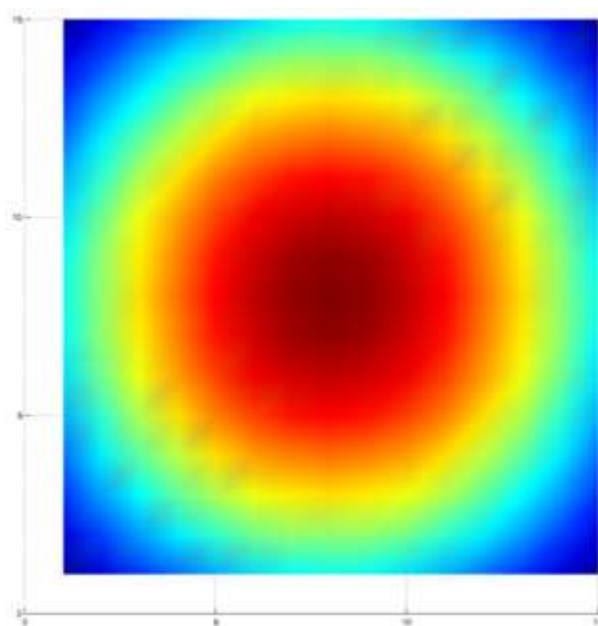
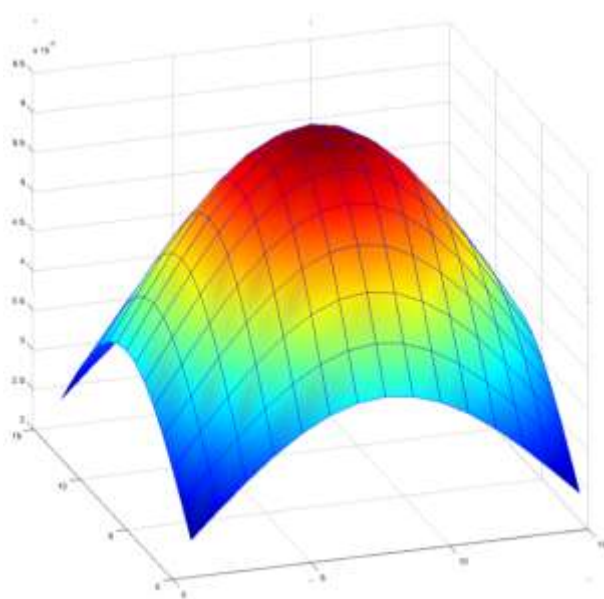
- [17]. **PAN, Ch.; LU, H.; CAO, F.** Segmentation bone marrow cell images via learning by sampling. [editor] HUANG. D. S. *ICIC*. místo neznámé: Springer-Berlin Heidelberg, 2009. stránky 336-345.
- [18]. **MAO-JUN, S.; ZHAO-BIN, W.; HONG-JUAN, Z.; M., YI-DE.** A new method for blood cell image segmentation and counting based on pcnn and autowave. *International Symposium on. Communications, Control and Signal Processing*. [konferenční publikace]. St Julians, Malta : IEEE Xplore, 12.-14.. 3 2008. stránky 6-9. E-ISBN: 978-1-4244-1688-2/08; DOI: 10.1109/ISCCSP.2008.4537182.
- [19]. **HE, W. a WILDER, J.** nucleus shape recognition for an automated hematology analyzing system. *Proceedings of the second joint EMBS/BMES Conference*. [sborník konference]. Houston: IEEE Xplore, 6. 1. 2002. 2. vyd., stránky 1043-1044. ISBN: 0-7803-7612-9/02; DOI: 10.1109/IEMBS.2002.1106267.
- [20]. **SABERI, A. a KHEIRI, K.** Granulometry Deficiencies in Blood Smears Particles Specification. *International Conference on Signal and Image processing*. [sborník konference]. Chennai: IEEE Xplore, 15-17. 2. 2010. stránky 271-274. ISBN: 978-1-4244-8594-9/10; DOI: 10.1109/ICSIP.2010.5697482.
- [21]. **JIANG, K., LIAO, Q. a DAI, S.** A novel white blood cell segmentation scheme using scale-space filtering and watershed clustering. Xian: 2003 International Conference on Machine Learning and Cybernetics, 2003. Sv. 5, 3, stránky 2820-2825. Dostupné z URL:
<http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1260033&isnumber=28160>.
DOI: 10.1109/ICMLC.2003.1260033.
- [22]. **REZATOFIGHI, S. H., a další.** A new aproach to white blood cell nucleus segmentation based on Gram-Schmidt orthogonalization. *Internatonal Conference on Digital Image Processing*. [sborník konference]. Bangkok: IEEE Xplore, 17-19. 3. 2009. stránky 107-111. ISBN: 978-0-7695-3565-4/09; DOI: 10.1109/ICDIP.2009.19.
- [23]. **KUMAR, R. D., JOSEPH a SREENIVAS, T. V.** Teager energy based blood cell segmentation. místo neznámé: IEEE XPLORE, 2002. stránky 619-622. DOI: 0-7803-7503-3/02.
- [24]. **BANDYOPAHYAY, S.** Method for cell segmentation. *Journal of Global Research in Computer Science*. 2011, Sv. 2., 4. vyd., stránky 130-135.

- [25]. **COLANTONIO, S., a další.** Cell Image Analysis Ontology. *Pattern Recognition and Image Analysis*. místo neznámé: Pleiades Publishing, Ltd., 2008. Sv. 18, 2. vyd., stránky 332-341. ISSN: 1054-6618; DOI: 10.1134/S1054661808020211.
- [26]. **BORAY TEK, F. G., DEMPSTER, A a KALE, I.** Blood cell segmentation using minimum area watershed and circle radon transformation. *Mathematical Morphology 40 years on*. Nietherlands: Springer, 2005. stránky 441-454.

Příloha 1 Charakteristiky Gaussových filtrů

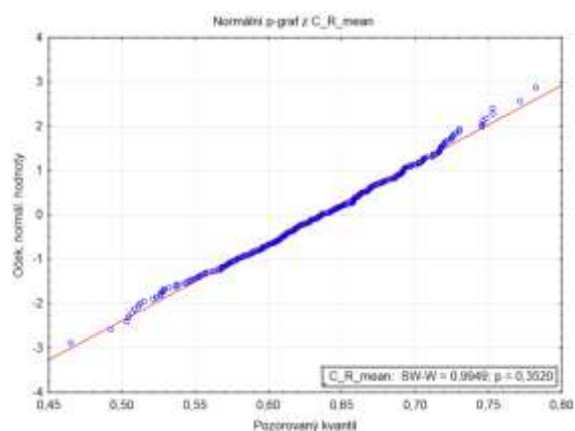
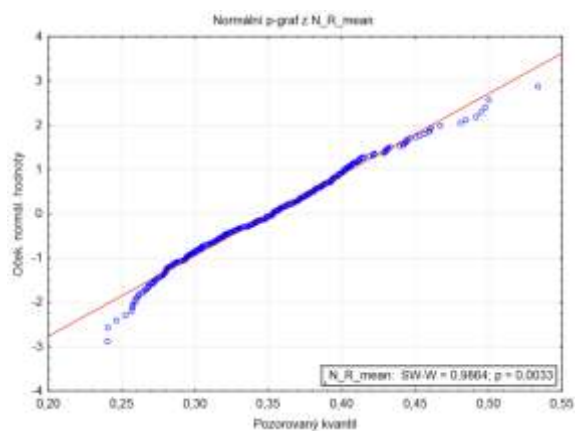


Gaussův filtr o velikosti 9x9 s parametrem sigma 5, využívaný před segmentací jádra

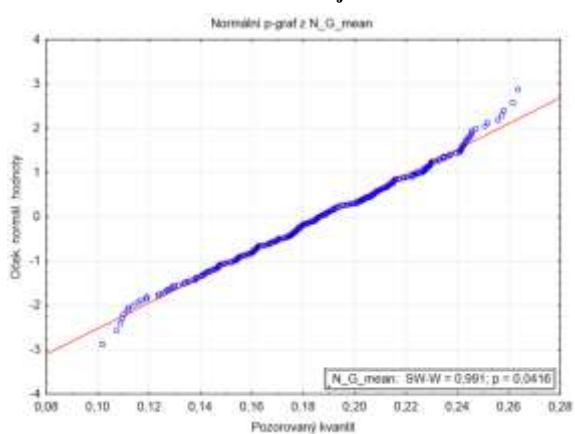


Gaussův filtr o velikosti 15x15 s parametrem sigma 7, využívaný před segmentací cytoplazmy

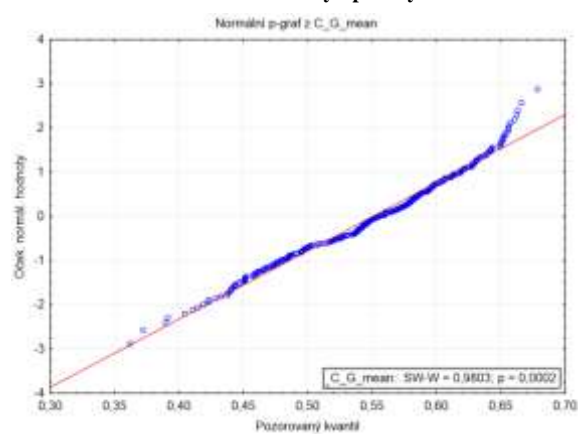
Příloha 2 Testy normality intenzitních příznaků jádra a cytoplazmy



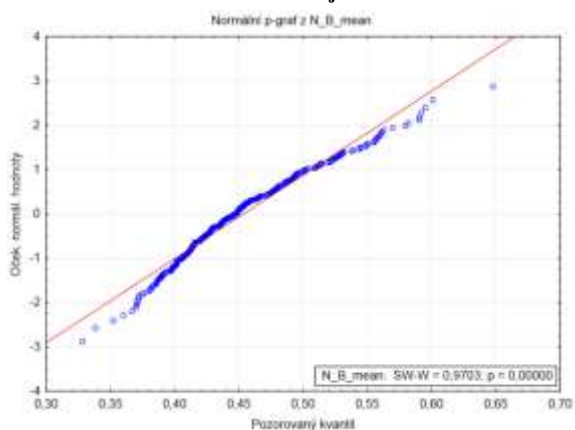
Červená složka jádra



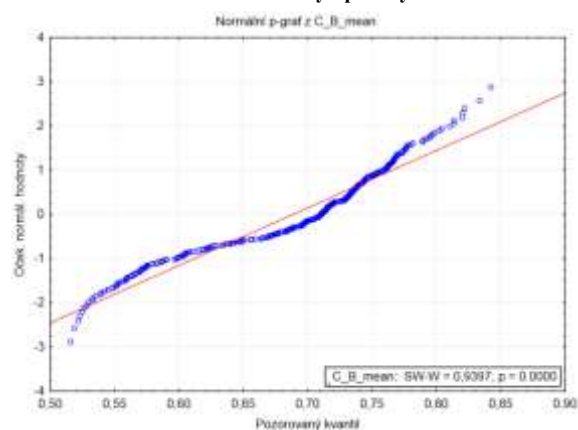
Červená složka cytoplazmy



Zelená složka jádra



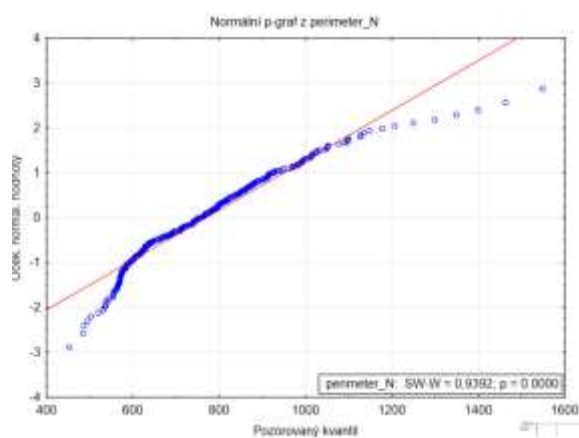
Zelená složka cytoplazmy



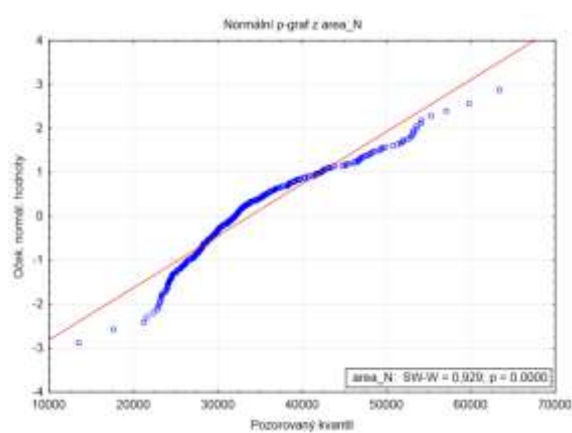
Modrá složka jádra

Modrá složka cytoplazmy

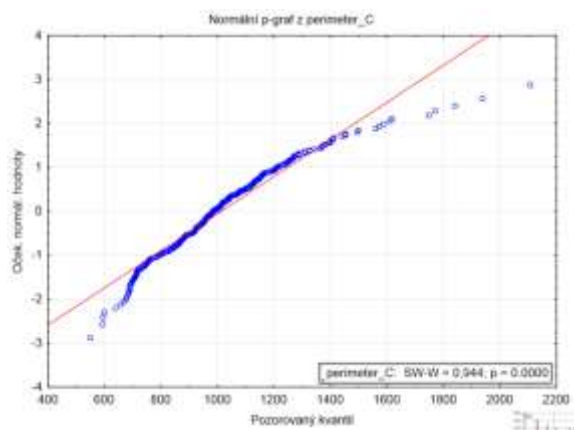
Příloha 3 Testy normality geometrických příznaků jádra a cytoplazmy



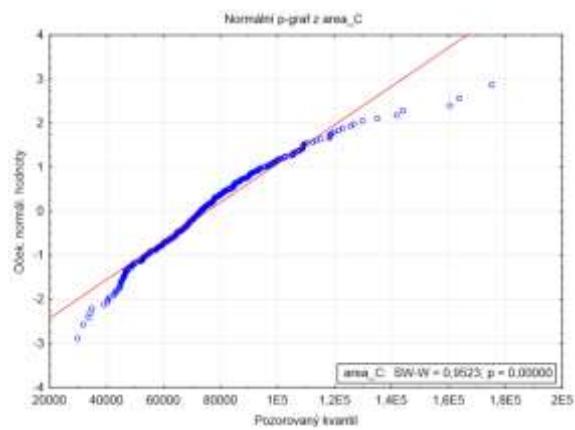
Obvod jádra



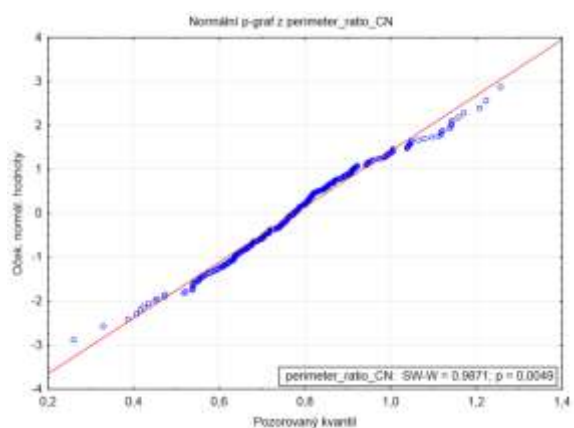
Plocha jádra



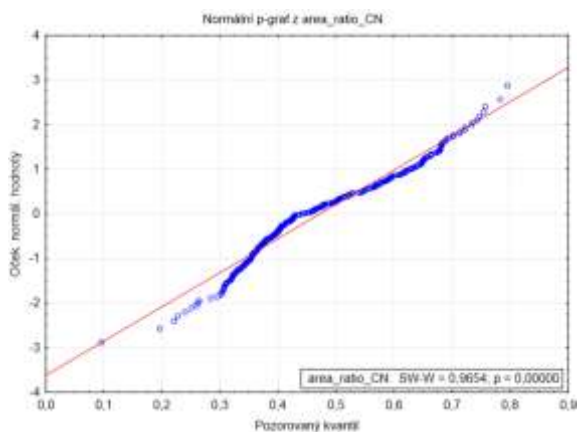
Obvod cytoplazmy



Plocha cytoplazmy

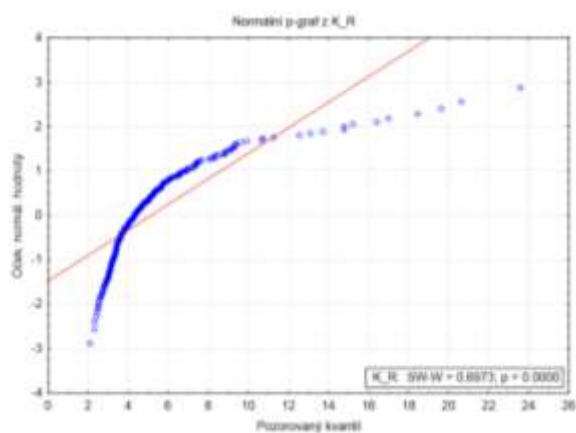


Poměr obvodů cytoplazmy a jádra

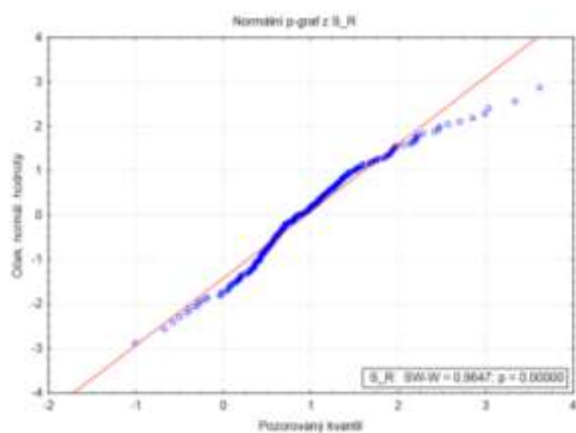


Poměr plochy cytoplazmy a jádra

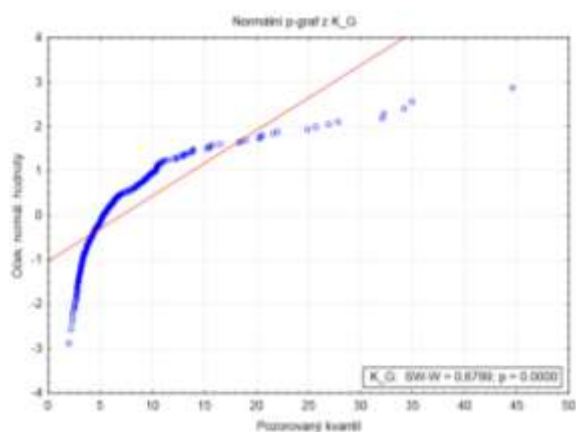
Příloha 4 Testy normality příznaků šikmosti a špičatosti jádra a cytoplazmy



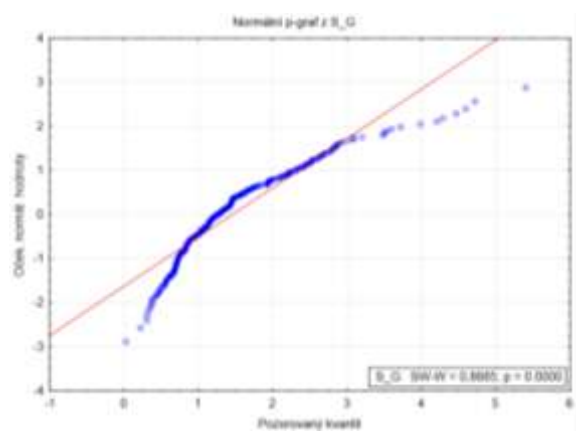
Šikmost červené složky histogramu



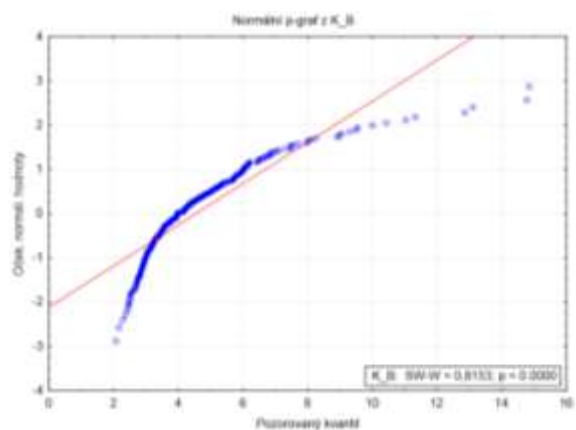
Špičatost červené složky histogramu



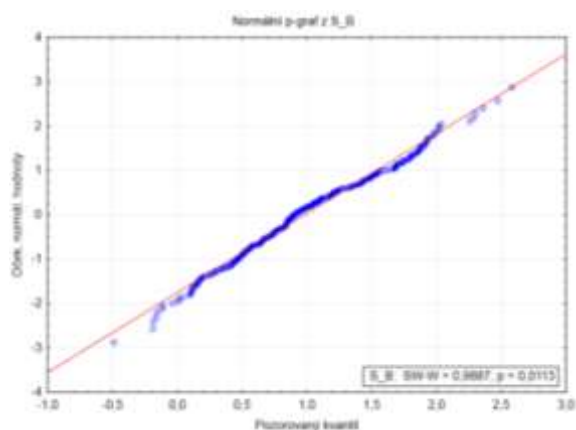
Šikmost zelené složky histogramu



Špičatost zelené složky histogramu



Šikmost modré složky histogramu



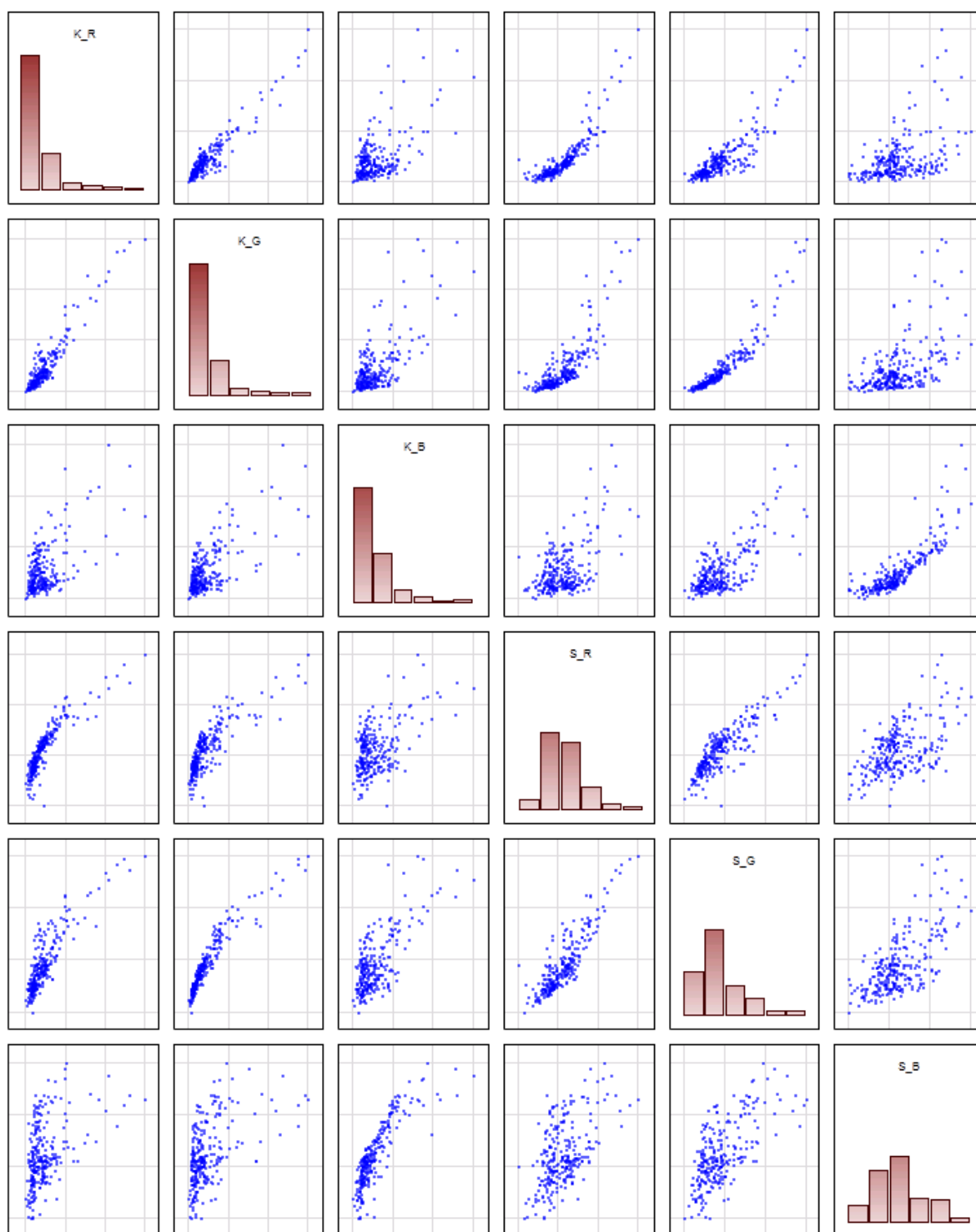
Špičatost modré složky histogramu

Příloha 5 Korelační grafy příznaků šikmosti a špičatosti histogramu cytoplazmy

K – kurtosis – šikmost histogramu cytoplazmy

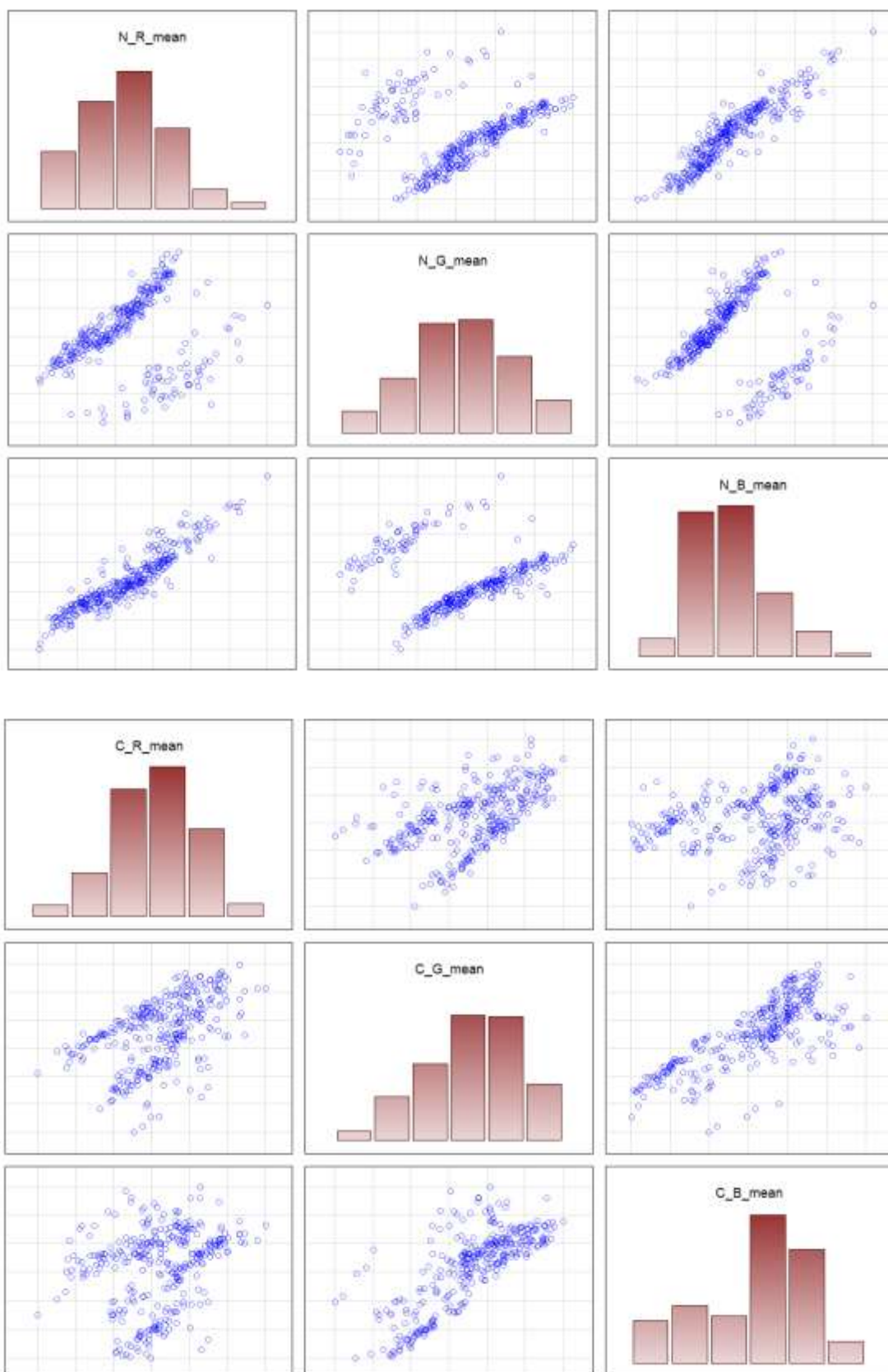
S – skewness – špičatost histogramu cytoplazmy

R, G, B – jednotlivé barevné složky histogramu



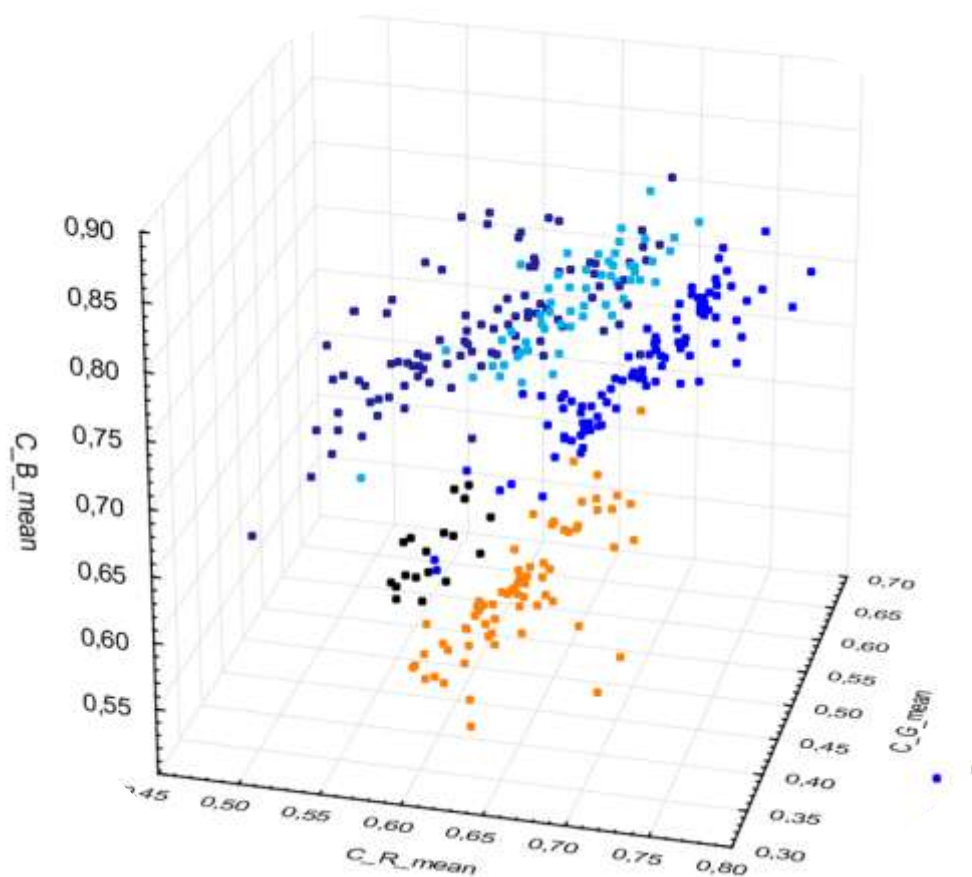
Příloha 6 Korelační grafy intenzitních příznaků jádra a cytoplazmy

N – nukleus – jádro C – Cytoplazma R, G, B – průměr intenzit cytoplazmy



Příloha 7 Graf intenzitních příznaků cytoplazmy před korekcí

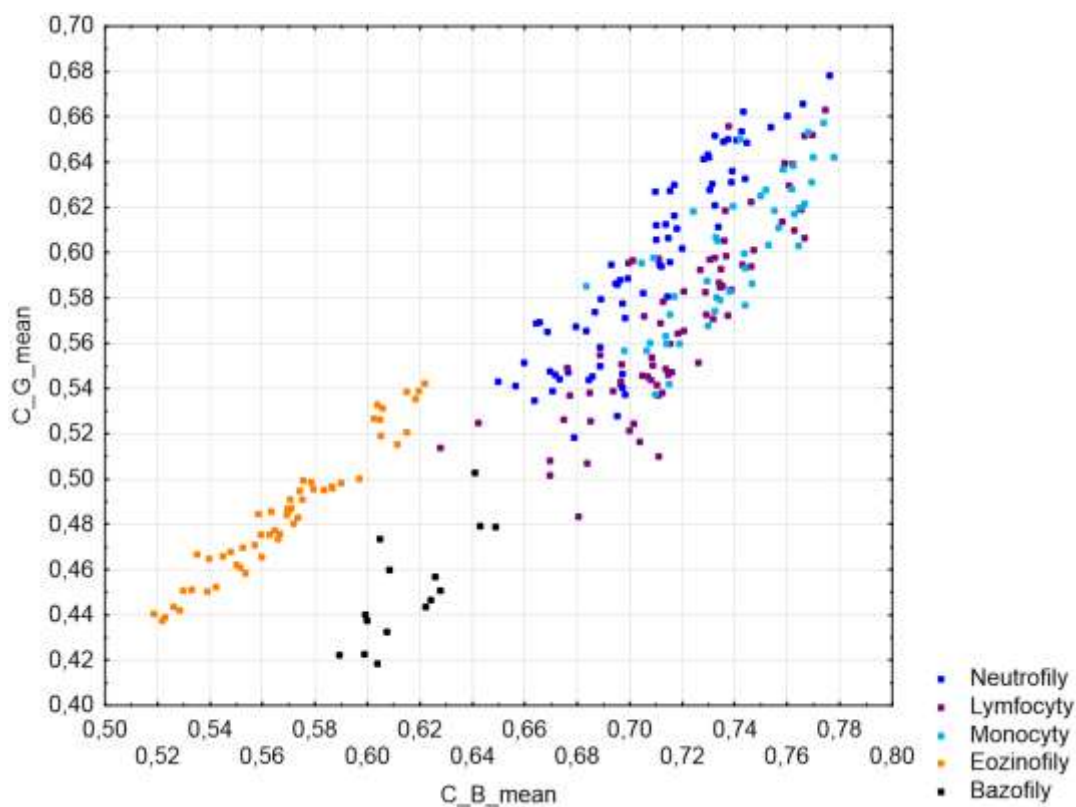
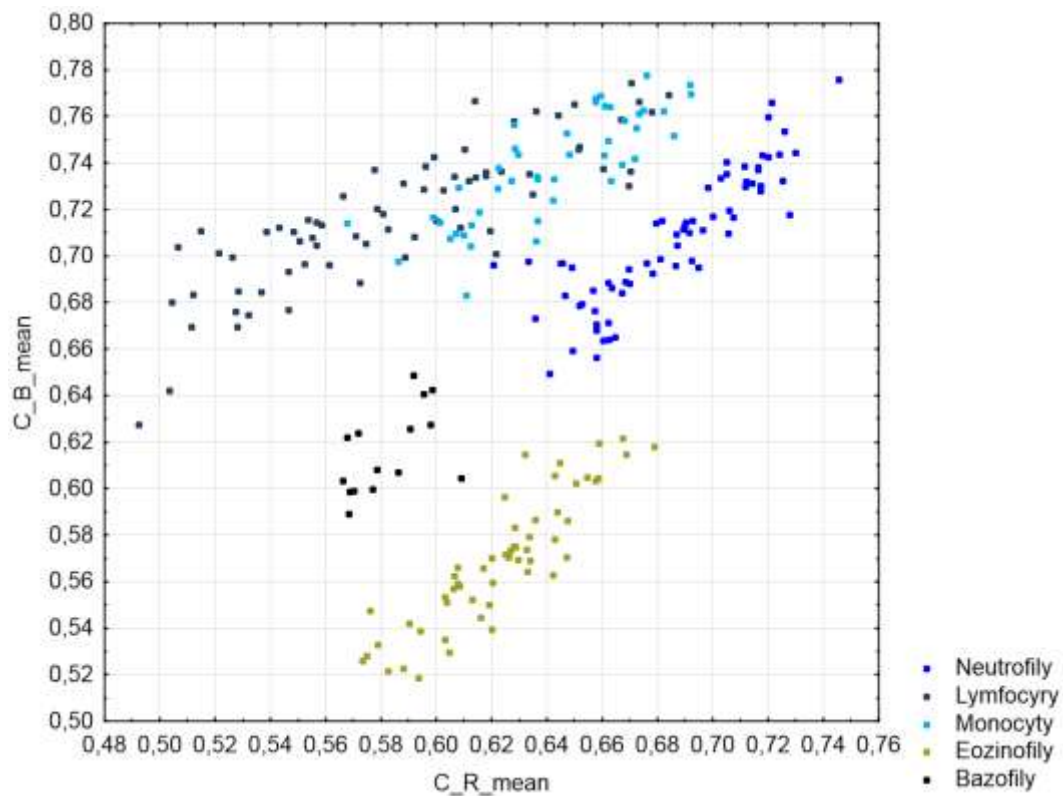
Tento graf zobrazuje původní množinu všech nasnímaných buněk, shluky jsou více rozptýlené a jejich kraje zasahují do jiných populací buněk. Vzdálené a nekonzistentní buňky byly vyřazeny z důvodu barevné odchylky.



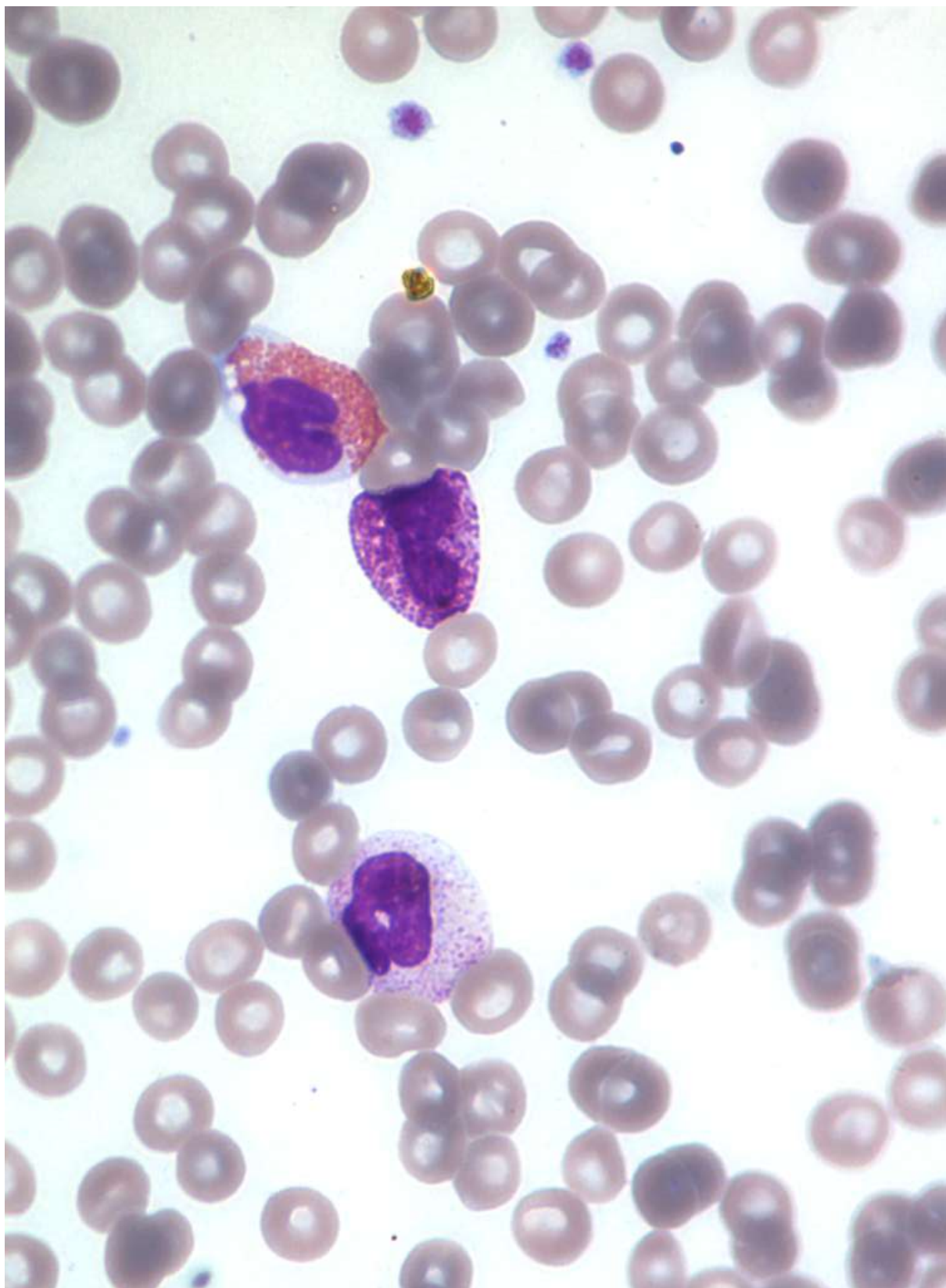
Příloha 8 Vybrané 2D grafy intenzitních příznaků cytoplazmy

C – cytoplazma

R, G, B průměrná intenzita RGB složek



Příloha 9 Případ výskytu více buněk v obraze



Případ výskytu více buněk v zorném poli mikroskopu, obraz je v původním rozlišení